Universitat de Girona

INTESTINAL MICROBIOLOGY IN CROHN'S DISEASE: A STUDY OF *Escherichia coli* AS A POTENTIAL ETIOLOGIC AGENT

Margarita MARTÍNEZ MEDINA

ISBN: 978-84-693-1991-8
Dipòsit legal: GI-324-2010
http://www.tdx.cat/TDX-0303110-152106

Universitat de Girona

TESI DOCTORAL

Intestinal microbiology in Crohn's disease:
A study of *Escherichia coli* as a potential etiologic agent.

Margarita Martínez Medina • 2009

PROGRAMA DE DOCTORAT

Química i Física de les Molècules i els
Materials, Biotecnologia i Ciències de la Salut

DIRIGIDA PER:

Prof. Dr. L. J. Garcia-Gil

Memòria presentada per a optar al títol de Doctor/a per la Universitat de Girona

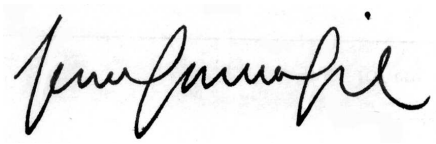
Universitat de Girona

Jesús Garcia Gil, catedràtic del Departament de Biologia de la Universitat de Girona,

CERTIFICO:

Que aquest treball titulat "Intestinal microbiology in Crohn's disease: A study of *Escherichia coli* as a potential etiologic agent", que presenta Margarita Martínez Medina per a l'obtenció del títol de Doctora, ha estat realitzat sota la meva direcció i que compleix els requeriments per poder optar a Menció Europea.

Signatura



Girona, Octubre de 2009

Universitat de Girona
Àrea de Microbiologia

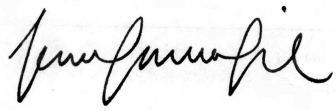
Tesi Doctoral

Intestinal microbiology in Crohn's disease: A study of *Escherichia coli* as a potential etiologic agent.

Memòria redactada per Margarita Martínez Medina, inscrita al programa de doctorat en Ciències: Química i Física de les Molècules i els Materials, Biotecnologia i Ciències de la Salut de la Universitat de Girona per optar al grau de Doctor en Biologia per la Universitat de Girona.

Vist-i-plau
El Director de la Tesi

Autor



Prof. Dr. L. J. Garcia-Gil
Catedràtic de Microbiologia

Margarita Martínez Medina

Girona, Octubre de 2009

Table of contents

Table of abbreviations

Abbreviation	Description	Abbreviation	Description
APCs	Antigen presenting cells	LPS	Lipopolysaccharide
APEC	Avian pathogenic *E. coli*	LRR	Leucine-rich repeats
ASCA	Anti-*Saccharomyces cerevisiae* antibody	MAP	*Mycobacterium avium* subspecies *paratuberculosis*
C	Control subjects	MAPK	Mitogen-activated protein kinase
CARD	Caspase recruitment domain	MDP	Muramyl-dipeptide
C-CD	Crohn's colitis	MHC	Major histocompatibility complex
CD	Crohn's disease	MNEC	Meningitis/sepsis-associated *E. coli*
CEACAM6	Carcinoembryonic antigen-related cell adhesion molecule 6	NF-kB	Nuclear factor-kappaB
DAEC	Diffusely adherent *E. coli*	NK	Natural killers
DC	Dendritic cells	NLR	NOD-like receptors
DEC	Diarrhoeagenic *E. coli*	NOD	Nucleotide-binding-oligomerisation-domain
DGGE	Denaturing gradient gel electrophoresis	Omp	Outer membrane protein
dsDNA	Double-stranded DNA	OTU	Operational taxonomic unit
EAEC	Enteroaggregative *E. coli*	PAI	Pathogenicity island
EHEC	Enterohemorrhagic *E. coli*	PAMPs	Pathogen associated molecular patterns
EIEC	Enteroinvasive *E. coli*	pANCA	Perinuclear anti-neutrophil cytoplasmic antibody
EPEC	Enteropathogenic *E. coli*	PGN	Peptidoglycan
ETEC	Enterotoxigenic *E. coli*	PRRs	Pattern recognition receptors
ExPEC	Extraintestinal pathogenic *E. coli*	RTi-PCR	Real time- polymerase chain reaction
FISH	Fluorescence *in situ* hybridisation	SCFAs	Short-chain fatty acids
GALT	Gut associated lymphoid tissue	SNPs	Single nucleotide polymorphisms
GI	Gastrointestinal	SSCP	Single strand conformation polymorphism
IBD	Inflammatory bowel diseases	(T)TGGE	(Temporal) temperature gradient gel electrophoresis
IBS	Irritable bowel syndrome	TIR	Toll-interleukin 1 receptor
IC	Ischemic colitis	TLR	Toll-like receptors
IC-CD	Ileo-colonic disease	TNF-α	Tumour necrosis factor alpha
I-CD	Crohn's ileitis	T-RFLP	Terminal- restriction fragment length polymorphism
IdC	Indeterminate colitis	UC	Ulcerative colitis
IFN	Interferon	UPEC	Uropathogenic *E. coli*
IL	Interleukin	UTI	Urinary tract infection

Summary

Crohn's disease is a chronic inflammatory bowel disorder of low mortality but high morbidity that afflicts both genders and any age group. Clinical symptoms are mainly diarrhoea, abdominal pain, and weight loss. However, multiple complications and extraintestinal manifestations are also usual. Industrialised countries have the highest incidence*[1] and prevalence* rates. Two-thousand new cases are diagnosed in Spain each year[2]. The etiology of Crohn's disease remains uncertain but recent progress has underscored the significance of genetic and immunologic features that confer susceptibility on the host, and external or environmental factors such as microorganisms and lifestyle.

The main purpose of this work was to describe the bacterial populations particularly related with Crohn's disease patients in order to identify putative etiologic agents (Chapter 1). To achieve this objective we analysed the overall microbial community associated with intestinal mucosa using a culture-independent, molecular-based approach. This first study offered us an initial idea about which bacterial populations were interesting and worth further study. Among them, *Escherichia coli* was more frequently found in Crohn's disease patients than control subjects. Although *E. coli* is known to be a common intestinal colonizer, previous research had already pointed to this bacterium as a putative etiologic agent. Moreover, a newly described pathovar named adherent-invasive *E. coli* (AIEC) has been recently linked to Crohn's disease. For that reason, the following studies were fucused on the *E. coli* mucosa-associated populations.

The second part of the study (Chapter 2) was aimed at describing the richness, abundance, diversity and pathogenic features of *E. coli*, particularly adherent-invasive strains, colonizing the intestinal mucosa. Approximately 100 *E. coli* colonies were isolated from ileal and colonic mucosa of 20 Crohn's disease patients and 28 controls. Two fingerprinting techniques, including pulsed field gel electrophoresis (PFGE), were used to analyse the clonality. Overall, the AIEC pathovar was searched for among 4314 isolates. The serotype, phylogenetic origin, and genotype (19 virulence genes) of *E. coli* and AIEC strains were also investigated. Although the number of different *E. coli* subtypes per patient was similar between Crohn's disease patients and controls, higher *E. coli* counts were characteristic of Crohn's disease patients (P=0.010), particularly those with Crohn's ileitis (P=0.001). Host-specific pulsotypes shared virulence features of extraintestinal pathogenic *E. coli* (ExPEC) at similar frequencies between

[1] Words marked with an asterisk in the text are defined in the Glossary (Annex III)
[2] Official webpage of the Asociación Española de Gastroenterología (http://www.aegastro.es)

Crohn's disease and controls. In contrast, higher AIEC prevalence (% subjects with AIEC: CD[3]= 51.9%; C[4]= 16.7%; P=0.003), abundance (% AIEC/*E. coli*: CD= 3.8 ± 5.0%; C= 1.5 ± 3.8%; P=0.039), and richness (Nº of AIEC subtypes: CD= 0.8 ± 1.4; C= 0.2 ± 0.4; P=0.015) were observed for Crohn's disease patients. In addition, AIEC subtypes showed a high variability of seropathotypes* and pulsotypes*, although the B2 phylogroup was the most prevalent (AIEC: 64%, non-AIEC: 38%, P=0.044). This is the fifth work since 2004 to describe the prevalence of AIEC among Crohn's disease patients. This exhaustive methodological approach has led to more accurate prevalence rates as well as to new information about the ecological parameters of the AIEC pathovar. It is the first study to depict the relative abundance, richness and diversity of AIEC strains within the mucosa-associated *E. coli* population. These new AIEC data reinforce it's implication in Crohn's disease.

Further studies aimed at characterising AIEC strains were performed on the collection of AIEC and non-AIEC strains obtained during the study mentioned above, which are presented in chapter 3.

Because bacterial biofilms in the gut mucosa are suspected to play a role in inflammatory bowel diseases, the main aim of chapter 3.1 was to compare the biofilm formation capacity of AIEC (N=27) and non-AIEC (N=38) strains isolated from the intestinal mucosa. Biofilm formation capacity was then contrasted with the AIEC phenotype[5], the serotype, the phylogenetic group and the presence of virulence genes. Specific biofilm formation indices were higher among AIEC than non-AIEC strains (P=0.007). In addition, 65.4% of moderate to strong biofilm producers were AIEC, whereas 74.4% of weak biofilm producers were non-AIEC (P=0.002). These data indicate that AIEC strains were more efficient biofilm producers than non-AIEC strains. Moreover, adhesion (P=0.009) and invasion (P=0.003) indices correlated positively with higher biofilm formation indices. Additionally, motility (100%, P<0.001), H1 flagellin (53.8%, P<0.001), serogroups O83 (19.2%, P=0.008) and O22 (26.9%, P=0.001), the presence of virulence genes such as *sfa/focDE* (38.5%, P=0.003) and *ibeA* (26.9%, P=0.017), and B2 phylogroup (80.8%, P<0.001) were frequent characteristics among biofilm producers. Biofilm formation capacity is a novel, complementary pathogenic feature of the recently described AIEC pathovar that could be implicated in Crohn's disease pathogenesis, by conferring to the pathovar a more perdurable colonisation of the intestinal tract, as well as protection against antimicrobial agents, thus leading to chronic infection.

[3] CD: Crohn's disease patients

[4] C: Control subjects (without inflammatory bowel disease)

[5] To date, AIEC straits can be only identified by phenotypic traits (adhesion to and invasion of intestinal epithelial cells, and intra-macrophage survival and/or replication)

Since several recent studies have pointed to a close relationship between AIEC and ExPEC in relation to the virulence genes that they carry in common, the main aim of chapter 3.2 was to determine the frequency of strains with 'AIEC phenotype' among a collection of ExPEC and further search for a common phylogenetic origin between those intestinal and extraintestinal AIEC strains. Adhesion, invasion and intra-macrophage replication abilities of 63 ExPEC strains were determined by a gentamicin protection assay using I407 and J774 cell cultures in order to identify their AIEC phenotype. Sixty-three ExPEC strains and 23 additional intestinal AIEC were compared by their virulence gene sets (*papC*, *sfa/focDE*, *afa/draBC*, *fimH*, *fimAv*$_{MT78}$, *hlyA*, *cnf1*, *cdt*, *iucD*, *neuC* and *ibeA*). In addition, we searched for correlations with AIEC phenotype, intestinal/extraintestinal origin, serotype and phylogroup. Phylogenetic relationships between extraintestinal and intestinal AIEC strains were determined using multilocus sequence typing of seven housekeeping genes (*adk*, *fumC*, *gyrB*, *icd*, *mdh*, *purA* and *recA*) and PFGE. Surprisingly, only 4 (6.35%) ExPEC strains, belonging to serotypes O6:H1 (two strains), O83:H1 and O25:H4 were classified as AIEC and found to be genetically related with some intestinal AIEC strains of the same serotypes as revealed by multilocus sequence typing (ST73, ST135 and ST131 respectively). No particular virulence gene sets correlated with the intestinal/extraintestinal origin of the strains, or with their AIEC-phenotype, whereas they did with the serogroup. It is also worth noting the identification of two intestinal AIEC strains and one extraintestinal AIEC with serotype O25:H4 that belonged to the emerging and virulent clonal group ST131. With this study we have demonstrated that the majority of extraintestinal pathogenic *E. coli* did not behave as AIEC did, despite a similar distribution of serotypes, phylogenetic groups and virulence gene profiles, thus confirming that the AIEC pathovar is close to ExPEC but has virulence-specific features that to date have only been detectable phenotypically. Further investigation determining the genes involved in the AIEC phenotype is necessary.

The results of this work are in agreement with previous research on the entire microbial community associated with Crohn's disease, providing evidence of a situation of dysbiosis. They further support the hypothesis of a putative implication of AIEC in this chronic inflammatory bowel disease. In addition, this work contributes to the understanding of mucosa-associated *E. coli* populations by providing new data about their ecological parameters and pathogenic features. Finally, we have also contributed to a better description of the AIEC pathovar by contrasting it with close pathogenic *E. coli* strains and finding new pathogenic features.

Resum

La malaltia de Crohn és una malaltia inflamatòria intestinal crònica que pot afectar homes i dones de totes les edats. Es caracteritza per presentar una baixa mortalitat, però elevada morbiditat. Els símptomes principals són diarrea (sovint sanguinolenta), dolor abdominal i pèrdua de pes. A més, sovint presenta diverses complicacions i manifestacions extraintestinals. Els països desenvolupats són els que tenen taxes de prevalença[1] *i incidència* més elevades. Cada any a Espanya es diagnostiquen 2000 casos nous d'aquesta malaltia[2]. L'etiologia de la malaltia de Crohn és encara desconeguda, malgrat la recerca intensiva feta darrerament. Actualment es pensa que hi participen factors genètics i immunològics que confereixen una susceptibilitat a l'hoste, i factors externs o ambientals, com serien els microorganismes i/o l'estil de vida.*

L'objectiu principal d'aquest treball ha estat descriure les poblacions bacterianes associades especialment als malalts de Crohn, amb la intenció d'identificar possibles agents etiològics (Capítol 1). Per tal de dur a terme aquest objectiu es va començar analitzant la composició global de la comunitat bacteriana present en la mucosa intestinal utilitzant mètodes moleculars. Aquest primer estudi ens va permetre identificar quines poblacions s'associaven als malalts de Crohn i quines eren més comunes en els controls sans. Entre aquestes, Escherichia coli *es va trobar més freqüentment en malalts de Crohn que en individus control. Malgrat que* E. coli *és un colonitzador comú del tracte intestinal, estudis previs fets per altres investigadors ja proposaven aquest microorganisme com a possible agent etiològic de la malaltia de Crohn. Recentment, a més, s'ha descrit un nou patovar anomenat* adherent-invasive E. coli (AIEC) *particularment associat a aquesta malaltia. És per això que ens vàrem centrar en les poblacions d'*E. coli *associades a la mucosa intestinal en estudis posteriors.*

*L'objectiu principal de la segona part d'aquest treball (Capítol 2) ha estat descriure la riquesa, abundància, diversitat i caràcter patogènic de les poblacions d'*E. coli *i* AIEC *presents en la mucosa intestinal. Per a fer-ho es va dur a terme l'aïllament d'aprop de 100 colònies d'*E. coli *de la mucosa ileal i colònica de 20 malalts de Crohn i 28 controls. Es van aplicar dues tècniques per a l'anàlisi de clonalitat dels aïllats, incloent l'electroforesi de camp polsant* (PFGE). *La identificació de soques pertanyents al patovar* AIEC *es va dur a terme sobre 4314 aïllats. A més, es va determinar el serotip, filogrup i genotip (19 gens de virulència) dels diferents subtipus d'*E. coli *i* AIEC *obtinguts. Malgrat les similituds de riquesa i diversitat de subtipus present en malalts de Crohn i controls, l'abundància d'*E. coli *era superior en els pacients de Crohn,*

[1] *Les paraules marcades amb un asterisc en el text estan definides en el Glossari (Annex III)*
[2] *Web oficial de la Asociación Española de Gastroenterología* (http://www.aegastro.es)

*especialment en aquells amb afectació ileal (P=0.001). Els clons aïllats eren específics de cada hoste, la qual cosa indica l'absència d'un clon o grup clonal comú entre els malalts de Crohn. Les soques compartien gens de virulència característics d'*E. coli *causants de malalties extraintestinals amb una freqüència similar en malats de Crohn i controls. Per contra, la prevalença (% d'individus amb* AIEC: *CD*[3]*= 51.9%; C*[4]*= 16.7%; P=0.003), l'abundància (% d'*AIEC/E. coli: *CD= 3.8 ± 5.0%; C= 1.5 ± 3.8%; P=0.039) i la riquesa (nombre de subtipus d'*AIEC *per pacient: CD= 0.8 ± 1.4; C= 0.2 ± 0.4; P=0.015) d'*AIEC *era superior en malalts de Crohn. Les soques d'*AIEC *presentaren una gran variabilitat de serotips i genotips, però el filogrup B2 fou el més prevalent (*AIEC*: 64%, no-*AIEC*: 38%, P=0.044). Després del primer treball publicat l'any 2004, aquest és el cinquè treball que descriu la prevalença de soques* AIEC *en malalts de Crohn. L'exhaustiva aproximació metodològica emprada va permetre donar dades de prevalença més acurades i obtenir informació sobre paràmetres ecològics específics del patovar* AIEC *dins l'intestí, fins al moment no descrits. En general, les dades obtingudes en aquesta part del treball recolzen la hipòtesi que el patovar* AIEC *està implicat en la malaltia de Crohn.*

Al capítol 3 es presenten els posteriors treballs de caracterització de les soques AIEC *i no-*AIEC *obtingudes i presentades al capítol 2.*

Es pensa que els biofilms bacterians que es troben en la mucosa intestinal juguen un paper important en la malaltia de Crohn, per això l'objectiu del capítol 3.1 d'aquest treball ha estat determinar la capacitat de formar biofilms de les soques AIEC *i comparar-la amb soques no-*AIEC*. L'índex de formació de biofilms va ser contrastat amb, a més del fenotip* AIEC[5]*, el serotip, filogrup i gens de virulència de les soques. Com a resultat d'aquest estudi es va observar que les soques* AIEC *presentaven índexs de formació de biofilm superiors que les soques no-*AIEC *(P=0.007) i que el 65.7% de soques amb capacitat moderada-forta de formar biofilms eren* AIEC*. A més, els índexs d'adhesió (P=0.009) i invasió (P=0.003) es correlacionaven positivament amb la capacitat de formar biofilms. La motilitat (100%, P<0.001), el tipus de flagelina H1 (53.8%, P<0.001), els serogrups O83 (19.2%, P=0.008) i O22 (26.9%, P=0.001), la presència de gens de virulència com* sfa/focDE *(38.5%, P=0.003) i* ibeA *(26.9%, P=0.017), i el filogrup B2 (80.8%, P<0.001) eren característiques freqüents entre les soques formadores de biofilms. La principal contribució d'aquesta part de l'estudi ha estat descriure la capacitat de formar biofilms* in vitro *com a característica fenotípica associada al patovar* AIEC *que podria tenir implicacions en la patogènia d'aquest en la malaltia de Crohn, ja sigui conferint al patovar una colonització més estable de la mucosa, com conferint una protecció envers agents antimicrobians, que conjuntament podrien col·laborar al fet que la infecció esdevingui crònica.*

[3] *CD: Pacients amb malaltia de Crohn*

[4] *C: Individus control (sense malaltia inflamatòria intestinal)*

[5] *A data d'avui, la identificació de soques AIEC només es pot fer a partir de la determinació de certs caràcters fenotípics com son: la capacitat d'adhesió i invasió de cèl·lules intestinals i la capacitat de supervivència i/o replicació en macròfags.*

Donada la similitud observada pel que fa als gens de virulència entre el patovar AIEC *i altres* E. coli *patògenes causants de malalties extraintestinals* (ExPEC)*, l'objectiu principal del capítol 3.2 ha estat determinar la freqüència de soques* ExPEC *amb fenotip* AIEC*, i després cercar l'existència d'un possible origen filogenètic comú entre aquelles soques* AIEC *d'origen extraintestinal i intestinal. La capacitat d'adhesió, invasió i de replicació en macròfags de 63 soques* ExPEC *es va determinar mitjançant cultius* in vitro *de cèl·lules I407 i J774 per tal de determinar-ne el fenotip* AIEC*. També es va comparar la distribució de gens de virulència* (papC, sfa/focDE, afa/draBC, fimH, $fimAv_{MT78}$, hlyA, cnf1, cdt, iucD, neuC, *i* ibeA) *entre aquestes 63* ExPEC *(aïllades principalment de casos d'infeccions del tracte urinari, septicèmia i meningitis) amb 23* AIEC *intestinals addicionals, tot correlacionant els grups de gens de virulència amb el serotip, filogrup, origen i fenotip* AIEC *de les soques. Es van utilitzar dos mètodes per a determinar la relació genètica entre les soques* AIEC *extraintestinals i intestinals:* MLST (multilocus sequence typing) *i* PFGE*. Per al* MLST *es van seqüenciar 7 gens de conservació de proteïnes estructurals* (adk, fumC, gyrB, icd, mdh, purA, *i* recA)*. Cal destacar que només 4 (6.35%) soques* ExPEC*, amb serotips O6:H1 (dues soques), O83:H1 i O25:H4, presentaren fenotip* AIEC*. Es va determinar, però, una relació genètica entre aquestes i d'altres* AIEC *d'origen intestinal amb els mateixos serotips mitjançant* MLST *(ST73, ST135 i ST131, respectivament). Cap col·lecció particular de gens de virulència no es va relacionar amb l'origen intestinal o extraintestinal de les soques, ni amb el fenotip* AIEC*, en canvi, les soques sí que es segregaven en funció del serogrup. És remarcable la identificació de dues soques* AIEC *intestinals i una extraintestinal amb serogrup O25:H4 que corresponien al grup clonal virulent i emergent ST131. Aquest estudi ha permès demostrar que la majoria d'*ExPEC *no es comportaven com* AIEC *malgrat la seva similitud envers els gens de virulència que presenten i els filotips i serotips als quals pertanyen. Amb això, es confirma que el patovar* AIEC *és proper a* ExPEC*, però posseeix característiques específiques relacionades amb la seva virulència que, fins ara, només es poden determinar fenotípicament. Són necessaris estudis posteriors que tinguin com a objectiu clarificar quina és la maquinària genètica implicada a conferir el fenotip* AIEC*.*

Els resultats d'aquest treball coincideixen amb investigacions prèvies que descriuen l'alteració bacteriana present en els malalts de Crohn. També recolzen la hipòtesi que implica el patovar AIEC *en l'etiologia d'aquesta malaltia inflamatòria intestinal. A més, contribuïm a la descripció de les poblacions d'*E. coli *associades a la mucosa intestinal aportant dades sobre aspectes ecològics i patogènics. Finalment, descrivim nous aspectes fenotípics d'*AIEC *que podrien estar relacionats amb la seva patogènia.*

Resumen

La enfermedad de Crohn es una enfermedad inflamatoria intestinal crónica que puede afectar tanto a hombres como a mujeres de distintas edades. Presenta una baja mortalidad pero elevada morbilidad. Los síntomas principales son la presencia de diarrea, a menudo con sangre, dolor abdominal y pérdida de peso, además de múltiples complicaciones y manifestaciones extraintestinales. La prevalencia[11] *e incidencia* son mayores en países desarrollados. Cada año se diagnostican en España unos 2.000 casos nuevos de la enfermedad*[12]. *La etiología de la enfermedad de Crohn se desconoce todavía, a pesar de la intensa investigación realizada. Se han implicado factores genéticos e inmunológicos que confieren susceptibilidad al individuo, además de factores externos o ambientales, como serian los microorganismos y/o el estilo de vida.*

El objetivo principal del presente trabajo ha sido describir las poblaciones bacterianas asociadas especialmente a los enfermos de Crohn, con la intención de identificar posibles agentes etiológicos (Capítulo 1). Para ello, comenzamos analizando la composición global de la comunidad bacteriana presente en la mucosa intestinal utilizando métodos moleculares. Este primer estudio permitió identificar qué especies se hallaban de forma más prevalente en enfermos de Crohn que en individuos sanos. Escherichia coli *fue una de las especies bacterianas que se esoció a los enfermos de Crohn. A pesar de que* E. coli *es un microorganismo común del tracto intestinal, estudios previos realizados por otros investigadores ya apuntaban hacia este microorganismo como posible agente etiológico de la enfermedad de Crohn. Además, recientemente, se ha descrito un nuevo patovar denominado* adherent-invasive E. coli (AIEC) *que se ha asociado a dicha enfermedad. Por esta razón, nos hemos centrado en las poblaciones de* E. coli *asociadas a la mucosa intestinal en estudios posteriores.*

El objetivo principal de la segunda parte del trabajo (Capítulo 2) fue el de describir la riqueza, abundancia, diversidad y carácter patogénico de las poblaciones de E. coli *y* AIEC *presentes en la mucosa intestinal. Para llevar a cabo este objetivo, se realizó el aislamiento de alrededor de 100 colonias de* E. coli *de la mucosa ileal y colónica d 20 pacientes de Crohn y 28 controles. Se utilizaron dos técnicas para analizar la clonalidad de los aislados, entre ellas la electroforesis en campo pulsado* (PFGE). *La identificación de cepas pertenecientes al patovar* AIEC *se realizó sobre 4314 aislados. Además, se determinó el serotipo, filogrupo, y genotipo (19 genes de virulencia) de los distintos subtipos de* E. coli *y* AIEC *obtenidos. A pesar de las similitudes de*

[11] *Las palabras marcadas con un asterisco en el texto están definidas en el Glosario (Annex III)*
[12] *Web oficial de la Asociación Española de Gastroenterología* (http://www.aegastro.es)

riqueza y diversidad de subtipos presentes en enfermos de Crohn e individuos control, la abundancia de E. coli *era superior en enfermos de Crohn, especialmente en aquellos pacientes con afectación ileal (P=0.001). Se hallaron clones específicos en cada huesped, excluyendo la existencia de un clon o grupo clonal común de los enfermos de Crohn. Las cepas compartían genes de virulencia característicos de* E. coli *causantes de infecciones extraintestinales, con una frecuencia similar en enfermos de Crohn y controles. En cambio, la prevalencia (% de individuos con* AIEC: *CD*[13]*= 51.9%; C*[14]*= 16.7%; P=0.003), la abundancia (% de* AIEC/E. coli: *CD= 3.8 ± 5.0%; C= 1.5 ± 3.8%; P=0.039) y la riqueza (número de subtipos de* AIEC *por paciente: CD= 0.8 ± 1.4; C= 0.2 ± 0.4; P=0.015) de* AIEC *era superior en enfermos de Crohn. Las cepas* AIEC *presentaron una gran variabilidad de serotipos y genotipos, pero el filogrupo B2 fue el más abundante entre ellas (*AIEC*: 64%, no-*AIEC*: 38%, P=0.044). Este es el quinto trabajo que describe la prevalencia de cepas* AIEC *en enfermos de Crohn, después de que el primero se publicase el año 2004. La exhaustiva aproximación metodológica utilizada permitió obtener valores de prevalencia más precisos, a la vez de obtener información acerca de parámetros ecológicos específicos del patovar* AIEC*, hasta el momento, no descritos. En general, los datos obtenidos en esta parte del trabajo apoyan la hipótesis que el patovar* AIEC *esta implicado en la enfermedad de Crohn.*

En el capítulo 3 se presentan los posteriores estudios de caracterización de cepas AIEC *y no-*AIEC *obtenidas y presentadas en el capítulo 2.*

Los biofilms bacterianos que se encuentran en la mucosa intestinal se consideran importantes en la etiología y/o desarrollo de la enfermedad de Crohn. Por este motivo, el objetivo del capítulo 3.1 ha sido el de determinar la capacidad de formar biofilms de las cepas AIEC *y compararlas con cepas no-*AIEC*. El índice de formación de biofilms fue contrastado con, además del fenotipo* AIEC[15]*, el serotipo, el filogrupo y los genes de virulencia de las cepas. Fue interesante observar que las cepas* AIEC *presentaban índices de formación de biofilms superiores que las cepas no-*AIEC *(P=0.007) y que el 65.7% de cepas con una habilidad moderada-fuerte de formar biofilms eran* AIEC*. Además, los índices de adhesión (P=0.009) e invasión (P=0.003) se correlacionaban positivamente con la capacidad de formar biofilms. La motilidad (100%, P<0.001), el tipo de flagelina H1 (53.8%, P<0.001), los serogrupos O83 (19.2%, P=0.008) y O22 (26.9%, P=0.001), la presencia de genes de virulencia como* sfa/focDE *(38.5%, P=0.003) e* ibeA *(26.9%, P=0.017), y el filogrupo B2 (80.8%, P<0.001) eran*

[13] CD: *Pacientes con enfermedad de Crohn*

[14] C: *Individuos control (sin enfermedad inflamatoria intestinal)*

[15] *A día de hoy, la identificación de cepas AIEC solamente se puede realizar determinando ciertas características fenotípicas, que son la capacidad de adhesión e invasión de células epiteliales y la capacidad de supervivencia y/o replicación en macrófagos.*

características frecuentes entre las cepas formadores de biofilms. La principal contribución de esta parte del estudio ha sido describir la capacidad de formar biofilms in vitro como característica fenotípica asociada al patovar AIEC que podría tener implicaciones en la patogénesis de dicho patovar en la enfermedad de Crohn, ya sea confiriendo al patovar una colonización más estable de la mucosa, como confiriendo una protección contra agentes antimicrobianos, que conjuntamente podrían colaborar a que la infección sea crónica.

Dada la similitud observada en cuanto a los genes de virulencia entre el patovar AIEC y otras E. coli patógenas causantes de infecciones extraintestinales (ExPEC), el objetivo principal del capítulo 3.2 fue el de determinar la frecuencia de cepas ExPEC con fenotipo AIEC y luego buscar si existe un origen filogenético común entre las cepas AIEC de origen intestinal y extraintestinal. La capacidad de adhesión, invasión y replicación en macrófagos de 63 cepas ExPEC se determinó mediante cultivos in vitro con células I407 y J774 para determinar el fenotipo AIEC. También se comparó la distribución de genes de virulencia (papC, sfa/focDE, afa/draBC, fimH, $fimAv_{MT78}$, hlyA, cnf1, cdt, iucD, neuC, y ibeA) entre estas 63 cepas ExPEC (aisladas principalmente de casos de infección urinaria, sepsis y meningitis) y 23 cepas AIEC intestinales, correlacionando la colección de genes de virulencia con el serotipo, el filogrupo, el origen y fenotipo AIEC de las cepas. Se emplearon dos métodos para la determinar la relación genética entre las cepas AIEC intestinales y extraintestinales: MLST (multilocus sequence typing) y PFGE. Para el MLST se secuenciaron 7 genes de conservación de proteínas estructurales (adk, fumC, gyrB, icd, mdh, purA, y recA). Cabe destacar que solamente 4 (6.35%) cepas ExPEC, con serotipos O6:H1 (dos cepas), O83:H1 y O25:H4, presentaron fenotipo AIEC, pero se halló relación genética entre éstas y otras AIEC de origen intestinal con serotipos iguales mediante MLST (ST73, ST135 y ST131 respectivamente). Ninguna colección concreta de genes de virulencia se correlacionó con el origen intestinal o extraintestinal de las cepas, tampoco con el fenotipo AIEC. En cambio, las cepas sí se segregaban en función del serogrupo. Es remarcable la identificación de dos cepas AIEC intestinales y una AIEC extraintestinal con serogrupo O25:H4 que correspondían al grupo clonal emergente y virulento ST131. Este estudio ha permitido demostrar que la mayoría de ExPEC no se comportan como AIEC, a pesar de la similitud genética que existe entre ambos patovares. Por lo tanto, se confirma que el patovar AIEC es próximo al patovar ExPEC, pero posee características específicas relacionadas con su virulencia que, hasta el momento, solamente se pueden determinar fenotípicamente. Hacen falta estudios adicionales que tengan como objetivo identificar cuál es la maquinaria genética implicada en conferir el fenotipo AIEC.

Los resultados de este trabajo coinciden con investigaciones previas que describen la alteración bacteriana presente en enfermos de Crohn. Además, apoya la hipótesis que implica el patovar AIEC *como agente etiológico de dicha enfermedad inflamatoria intestinal. Contribuimos también en la descripción de las poblaciones de* E. coli *asociadas tanto a la mucosa intestinal de individuos sanos como pacientes de Crohn aportando datos acerca de aspectos ecológicos y patogénicos de éstas, así como en la caracterización de cepas* AIEC, *hallando nuevos aspectos fenotípicos que podrían estar relacionados con su patogénesis.*

Introduction

1 Inflammatory bowel diseases

Inflammatory bowel diseases (IBDs) include various different intestinal inflammatory disorders. A common feature is chronic inflammation of different parts of the gastrointestinal (GI) tract. Infectious colitis, ischemic colitis and radiation enterocolitis are disorders of known etiology but there are still some types of IBDs of unknown cause, called idiopathic IBDs. Crohn's disease (CD), ulcerative colitis (UC), and indeterminate colitis are the primary chronic idiopathic IBDs. Despite the fact these intestinal disorders differ in terms of their localisation, the distribution of inflamed areas, and their histology, in some cases a clear classification is not possible [1].

During the course of IBDs, disease-active periods, or **flares,** alternate with inactive episodes of **remission**. The macroscopic re-emergence of the disease for those patients who have undergone surgery is called **recurrence**. Three subgroups can be distinguished, based on the clinical course of the disease. IBDs are classified as **chronic-intermittent** when there is a relapsing remitting pattern in which flares respond favourably to treatment and temporary periods of remission are achieved. Cases of refractory disease are classified as **chronic-continuous**, because treatment only partially heals the inflammation, and remission periods are shorter and flares more frequent. The term used when the disease is accompanied by extremely severe inflammation is **acute-fulminant**. Commonly, this phase appears as the first episode of the disease, although it can emerge during both of the chronic types of the disease described above [1].

1.1 Crohn's disease

Inflamed areas in CD patients are patchily distributed and may involve the entire gut, from the oropharynx to the perianal area. The inflammation is distributed asymmetrically, not only along the digestive tract but also across the perimeter. It can be transmural*[16], thus affecting the whole intestinal wall from the mucosa to the serosa, which may lead to other complications such as fistulas*, abscesses* and stenosis*. Histological observations will reveal small superficial ulcerations and sometimes noncaseating granulomas* [2].

[16] Words marked with an asterisk in the text are defined in the Glossary (Annex III)

There are a great number of disease phenotypes, which can be classified in line with the Montreal classification [3] (Table 1) according to the age at the onset of the disease, the anatomical localisation of the inflammation, and its overall behaviour. Clinical manifestations are unstable during its course, and for this reason a long monitoring period (at least five years) is necessary to accurately classify the disease phenotype [4].

Table 1. The Montreal classification of Crohn's disease ([3]), modified from the Vienna classification ([5]).

Age of diagnose (A)	Characteristics
A1: Younger than 16	Colonic localisation in most cases High family aggregation and genetic susceptibility
A2: 17-40 years old	Frequent and extensive inflammation, from upper GI tract to colon
A3: Older than 40	Colonic localisation in most cases
Localisation (L)	**Characteristics**
L1: Ileal	30% CD patients Basic clinical manifestations: Stenosis*, nausea, vomiting, abdominal pain, loss of weight, and fever. Less aggressive diarrhoea than in colonic localisation.
L2: Colonic	20% CD patients One or several affected areas between cecum and rectum, but mainly colon. Abundant diarrhoea, bleeding, abdominal pain, and loss of weight. Correlates with perianal disease and extraintestinal manifestations.
L3:Ileocolonic	40% CD patients Localisation and clinical manifestations of L1 and L2
L4: Upper GI tract	5% CD patients Proximal ileum, jejunum, duodenum, stomach, oesophagus or oropharynx can be affected. Heterogeneous clinical manifestations depending on the exact localisation.
L4 Localisation must be added to the above categories in case of additional upper GI tract involvement (e.g. L1+L4)	
Behaviour (B)	**Characteristics**
B1: Inflammatory (not structuring-not penetrating)	Superficial ulcerations and inflammation Abdominal pain and diarrhoea
B2: Stricturing	Presence of stenosis* and fibrosis* Nausea, vomiting, pain and abdominal distension. Cases often refractory. Occasional surgical intervention. Low recurrence.
B3: Penetrating	Perforation. Often formation of fistulas* and abscesses*. Surgery necessary. High recurrence.
Perianal disease is a modulator of the above categories, which must be indicated with a *p* (e.g. B1p)	

1.2 Ulcerative colitis

UC is a form of IBD that occurs exclusively in the colon. The rectum is involved in 95% of patients, with variable degrees of proximal extension. Unlike CD, inflammation is confined to the mucosa except for rare cases, such as toxic megacolon*, and has a continuous distribution of variable severity. Ulceration, edema* and hemorrhage are also characteristic of UC. Histological features include acute and chronic inflammation by polymorphonuclear leukocytes and mononuclear cells, crypt abscesses*, distortion of the mucosal glands, and goblet cell depletion. The most common intestinal manifestation is the presence of blood and mucus mixed in feces, together with lower abdominal cramping [2]. Depending on the anatomical extension of the inflammation, the disease is subclassified as [3]:

- *Ulcerative proctitis* (E1) refers to inflammation that is limited to the rectum. Generally mild intermittent rectal bleeding is the only symptom. However, rectal pain, urgency and tenesmus* are also frequent. Prevalence* among UC patients is 30-42% [6].
- *Distal* or *left-sided UC* (E2) involves the rectum, sigma and left colon. Symptoms include bloody diarrhoea, abdominal cramps, weight loss and left-sided abdominal pain. Prevalence* among UC patients is 44-48% [6]. Some clinicians consider *proctosigmoiditis* to be a different subtype [1]. Nevertheless, disease course and prognosis are similar to left-sided UC.
- *Pancolitis* or *universal colitis* (E3) refers to inflammation affecting the entire colon. Bloody diarrhoea, abdominal pain and cramps, weight loss, fatigue, and fever are symptoms of pancolitis. Patients with this disease phenotype have a higher probability of colectomy and colorectal cancer. Prevalence* among UC patients is 9-17% [6]. Fulminant colitis is a rare but severe form of pancolitis. Patients with fulminant colitis show dehydration, severe abdominal pain, diarrhoea with bleeding and even shock. These patients have a risk of developing a toxic megacolon*.

1.3 Indeterminate colitis

The concept of indeterminate colitis (IdC) is controversial, since some clinicians consider that patients diagnosed with IdC are in fact suffering from either CD or UC that has not been properly identified, and that this classification is therefore temporary (10-15% of cases) [7]. Currently, however, it is generally considered to be a chronic IBD that principally affects the colon, and for a correct diagnosis infectious colitis or other causes of colitis must have been previously ruled out.

The lack of specific reference tests for this IBD phenotype make it difficult to diagnose, and can eventually lead to its misidentification as CD or UC. A lot of good judgement must be exercised in classifying unidentified IBDs because pharmacologic treatment and how and when surgery is used depends on it. The main clinical manifestation is severe colitis, with painful bloody diarrhoea and fever. Usually, patients suffering from IdC do not improve with corticosteroid treatment. They are generally refractory and surgery is usually necessary. IdC is characterised by extensive ulceration, alternating with non-inflamed areas (50%). It occurs at higher rates in the rectum and left colon. Under microscopic examination, this IBD exhibits abundant ulcers, V-shaped fistulas*, transmural* inflammation and an absence of lymphoid aggregates [8].

1.4 Extraintestinal manifestations of IBDs

Multiple organs can be affected by IBD, including bones, joints, skin, eyes, hepatobiliary systems, lungs and kidneys [9]. These extraintestinal manifestations can occur prior to the disease's onset, in conjunction with the disease or as a consequence of flares. The prevalence* of patients that suffer at least one extraintestinal manifestation ranges from 21 to 40% [10].

Arthralgia* and arthritis* [11]. Articular pathologies are the most frequent extraintestinal manifestations of IBD, with a prevalence* of 30% in all IBD patients, and a higher percentage in CD patients. There are two forms of musculoskeletal syndrome: a peripheral form and an axial form. Axial arthropathies, including sacroiliitis* and ankylosing spondylitis*, are usually independent of disease activity and involve inflammation of the spine and sacroiliac joints respectively. Peripheral arthritis* is typically classified in two forms: Type I and II. Type I involves less than five large joints, is acute, and usually correlates with flares. Type II tends to be chronic, affects five or more small joints, and is not strictly associated with IBD activity periods. Included within IBD musculoskeletal pathologies is osteoporosis.

Dermatologic. Oral aphthous ulcers are common in IBD and usually correlate with activity periods of the disease. Other cutaneous manifestations are erythema nodosum* and pyoderma gangrenosum*. Dermatologic pathologies are present in <8% of IBD patients, with erythema nodosum* being more frequent in CD patients and pyoderma gangrenosum* in UC patients (<1%) [2].

In addition to the above-mentioned extraintestinal manifestations, less frequent complications can sometimes appear. **Ophthalmologic** problems such as uveitis*, scleritis*, and episcleritis*, are often a secondary effect of chronic

corticosteroid therapy. **Hepatobiliary** (up to 30% of IBD patients, more common in UC), **vascular** and **renal diseases** (up to 16% of IBD patients), as well as **bone abnormalities** are other reported forms of IBD extraintestinal manifestations [1, 2, 9].

The wide range of extraintestinal manifestations related to IBDs may have either an immunologic etiology, dependent or independent of the disease activity, or occur due to anatomical or metabolic abnormalities of IBDs or as the secondary effects of drug treatment [12].

1.5 Other intestinal inflammatory diseases

Ischemic colitis [13]. Ischemic colitis (IC) is caused by an inadequate blood flow in the colon which leads to an acute inflammation. IC is more common in the elderly than in the young. Moderate abdominal pain and, rarely, hemorrhaging, are symptoms of IC. Histologically, different degrees of ischemia* can appear in both the mucosa and submucosa. In addition, transmural* necrotic zones can be present.

IC consists of an acute colonic inflammation that is distinguishable and has a different etiology from IBDs. For this reason, it is considered an excellent "inflammatory control" since it allows cases of patients (in this study, primarily CD patients) to be compared with a case of acute mucosal inflammation, arising out of a formerly supposedly healthy situation.

Irritable bowel syndrome. IBS is considered a functional bowel disorder [14], since clinical symptoms of gastrointestinal dysfunction are present, but clear endoscopic and histological evidence is absent. IBS is the most prevalent gastrointestinal disorder, with 15% of the population affected in Western countries [15]. Patients suffering from IBS may have abdominal pain and/or discomfort, bloating and bowel disturbances (diarrhoea, constipation or alternating bowel habits) [16]. IBS is subdivided into three groups according to the clinical symptoms: D-IBS, characterised by predominant diarrhoea; C-IBS, by predominant constipation, and A-IBS, by the subtypes alternating [17]. Close points of similarity between D-IBS and CD have lead to the hypothesis that IBS is an inflammatory disease that shares common pathogenic features with CD, but has a milder phenotype [18]. IBS is a complex disorder; the influence of possible alterations in the central and enteric nervous systems [19], as well as impaired permeability possibly triggered by a luminal antigen [20], have been considered as possible causes of or contributors to the disease.

2 Crohn's disease

2.1 Clinical manifestations and diagnosis

Crohn's disease is a chronic inflammatory disorder that can be patchily and asymmetrically localised throughout the entire gastrointestinal tract, longitudinally (from mouth to anus) and transversally (from mucosa to serosa). Factors such as localisation of the disease, severity, and presence of extraintestinal manifestations determine its clinical manifestation in patients. Symptoms are unspecific and heterogeneous, which makes diagnosis difficult. Various different diagnostic tools such as histology, image-based tools (endoscopy, radiology), and other complementary laboratory tests must be used.

The classic symptoms are diarrhoea, abdominal pain and weight loss. **Diarrhoea** is the most frequent symptom when the disease is first diagnosed. Rectorrhagia can appear in patients with L2 (colonic) localisation but less frequently than in UC patients. Tenesmus*, urgency and incontinence usually appear in those patients with rectal involvement. However, these symptoms are more typical of UC patients. The type and intensity of **abdominal pain** can vary depending on its nature and localisation. It is very frequent in paediatric patients (72%) [21]. **Loss of weight** is more characteristic of CD than UC. It is a consequence of the inflammatory process per se, of increasing catabolism and malabsorption, and also frequently of anorexia. **Perianal disease** (fistulas* and/or abscesses*) are frequent in CD patients with colonic localisation (40%), although 20% of all phenotypes of CD patients suffer from this complication during the disease's development ([22] and references therein). Another classic symptom of CD is **fever**; however, this usually comes out in combination with other symptoms such as abscesses* or perforation. A wide range of **extraintestinal manifestations,** summarised in Section 1.4, are additional complications that may occur in CD patients.

Clinical features, in conjunction with image-based tools and histology, are necessary for CD diagnosis. Colonoscopy with ileoscopy is a routine exploration when CD is suspected. **Endoscopic techniques** allow an overall description of which kind of lesions are present, in what amount, and any point where the disease is localised along the GI tract. In addition, biopsy sampling is possible during endoscopy. **Radiology** provides important additional data about the disease's behaviour (e.g. presence of fistulas*), which makes it useful for phenotype classification. Many **histological** features are used for IBD diagnosis [23], with the presence of granulomas* being a key characteristic of Crohn's

disease [24]. Other microscopic traits generally accepted as permitting a diagnosis of CD are the presence of focal and patchy chronic inflammation with increased lymphocytes and plasma cells infiltrates, and focal crypt irregularity (discontinuous crypt distortion), among others [23]. Several **biological markers** are used in combination with previously mentioned diagnostic techniques [25]. These tests comprise i) detection of serological markers such as perinuclear anti-neutrophilic cytoplasmic antibodies (pANCAs) and anti-*Saccharomyces cerevisiae* antibodies (ASCAs), ii) C-reactive protein quantification, iii) globular sedimentation rate determination, and iv) fecal calprotectin and lactoferrin concentrations. None of these biological markers is pathognomonic of CD; however they can be useful, in some cases, in the differential diagnosis of CD and UC. For example, ASCAs are more characteristic of CD than UC (50-60% of CD patients show positive results, in contrast with only 10% of UC patients), while the converse is true of pANCAs (70% of UC patients *vs* 10-40% of CD patients). In addition, some of them can be useful in classifying the disease phenotype and estimating its activity and progress (e.g. C-reactive protein quantification), as well as for prognostic purposes such as predicting the response to treatment. New serological markers have been already associated with IBDs (for a review on this subject, see [26]).

2.2 Etiology

Although the etiology of Crohn's disease remains elusive to date, it is widely accepted that several factors play a part in the onset and perpetuation of the disease. These factors include genetic and immunologic features that confer susceptibility on the host, and external or environmental factors such as microorganisms and lifestyle. Depending on the scope of their research, scientists differ in their view of the significance of each factor. Xavier and Podolsky (2007) [27] support the view that host factors implicated in barrier function and innate and adaptive immunity have a greater role than environmental factors. Nevertheless, they consider that microorganisms play a more important role than other environmental factors like lifestyle. In contrast, R.B. Sartor (2006) [28] gives equal importance to genetic susceptibility, immune response, luminal microbial antigens and other environmental triggers (Figure 1).

The evidence that associates these factors with the etiology of Crohn's disease are explained in separate sections. **Environmental factors** that can be a predisposition to CD are included in the next section (Section 2.3), which focuses on the epidemiologic features of CD. **Host factors**, including known **genetic** susceptibility loci and imbalances in **immune system** functions can be found in

Sections 2.5 and 3.3 respectively. Finally, **microbial influences** on CD are more extensively detailed in Section 4.

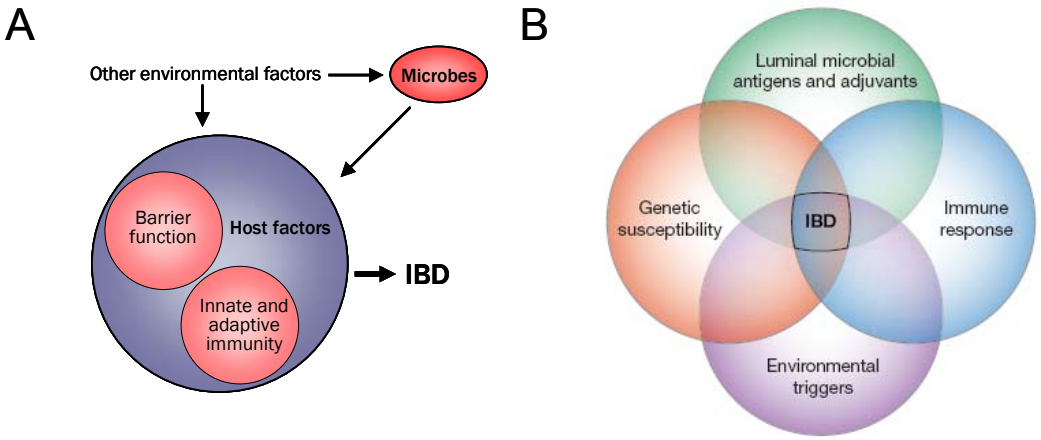


Figure 1. Etiological factors implicated in Crohn's disease onset or perpetuation. A: Model proposed by Xavier and Podolsky, 2007 [27]. B: Model proposed by R.B. Sartor, 2006 [28]

2.3 Epidemiology *(Environmental Factors)*

Higher rates of IBDs are seen in industrialised countries. In particular, the highest incidence* and prevalence* rates have been reported in Scandinavian countries, followed by the United Kingdom and North America, where there was an increase between the 1860s and 1980s, when a plateau was reached [29]. In Spain, as in other southern European countries, the incidence* is lower but is still increasing. In a retrospective study of 1966-1975 in Galicia, an incidence* of 0.1/100,000 per year was determined, while in a more recent prospective study (Asturias, 2000-2002) the incidence* had increased up to 7.5/100,000 per year ([30] and references therein). Prevalence* rates are higher than incidence* rates because of the chronic nature of the disease and the longevity of patients, reaching around 1% of the population in some areas. For a detailed review of the prevalence* and incidence* rates of studies carried out world-wide, see reference [30]. CD can occur in all ages and genders, but the peak age of onset is around 20 years old. The distribution of prevalence* among ages is bimodal, with a second peak of high prevalence* existing for people between 50 and 70 [31].

Regarding the existence of a clear north-south gradient, environmental factors and lifestyle seem to play a role in CD pathogenesis. Among these, changes in hygiene [32] and dietary habits [33], the use of refrigerators for food maintenance [34], and the increased use of antibiotics in human and veterinary medicine [35] might be contributing factors. Moreover, prevalence* rates in people that migrate

from undeveloped to industrialised countries equal those in the destination countries [29]. Racial differences are clear for Jewish people, who have higher incidence* rates than any other ethnic group, which is evidence of a possible genetic risk as well.

Many environmental factors have been associated with IBDs to date, but their role in the pathogenesis of the disease remains uncertain. A summary of IBD risk factors are listed in Table 2. Some of them behave differently in CD and UC, which suggests that they are heterogeneous disorders. In fact, there is an increasing tendency to believe that IBDs probably include more than two or three distinct entities.

Table 2. Summary of exogenous risk factors associated with ulcerative colitis and Crohn's disease. Adapted from Aladrén *et al.* [30] ([30, 31, 36] and references therein).

Risk Factor	UC	CD
Lactose intolerance	protective	protective
Tobacco (active smoking)	protective	risk factor
Appendectomy	protective	protective?
Natural feeding of newborns	protective?	protective?
Diet		
Refined carbohydrates	risk factor	risk factor
Fatty acids	risk factor?	risk factor?
Others (margarine, alcohol, etc.)	lack of evidence	lack of evidence
NSAIDs (Non-steroidal anti-inflammatory drugs)	controversial	controversial
Contraceptive hormones	controversial	controversial

2.4 Current therapies

The management of Crohn's disease depends on disease localisation (ileocecal/ colonic), severity (mild/ moderate/ severe), and activity (induction/ maintenance of remission) (Figure 2). The dose and supply method (oral, subcutaneous or intravenous) of a great range of drugs has to be chosen for each case. Common therapies currently used for the management of Crohn's disease are anti-inflammatory chemicals derived from salicylic acid, such as mesalazine and sulfasalazine; corticosteroids such as prednisone, methyl-prednisone and budesonide; antibiotics such as metronidazole and ciprofloxacin; immunosuppressors like azathioprine and mercaptopurine; antimetabolite and antifolate methotrexate; and, more recently, the so called "biological" drugs, consisting of antibodies against tumour necrosis factor α (TNF-α), such as infliximab and adalimumab. Surgery is appropriate for those patients with fulminant or fistulising disease and for refractory cases. For a review, see reference [36].

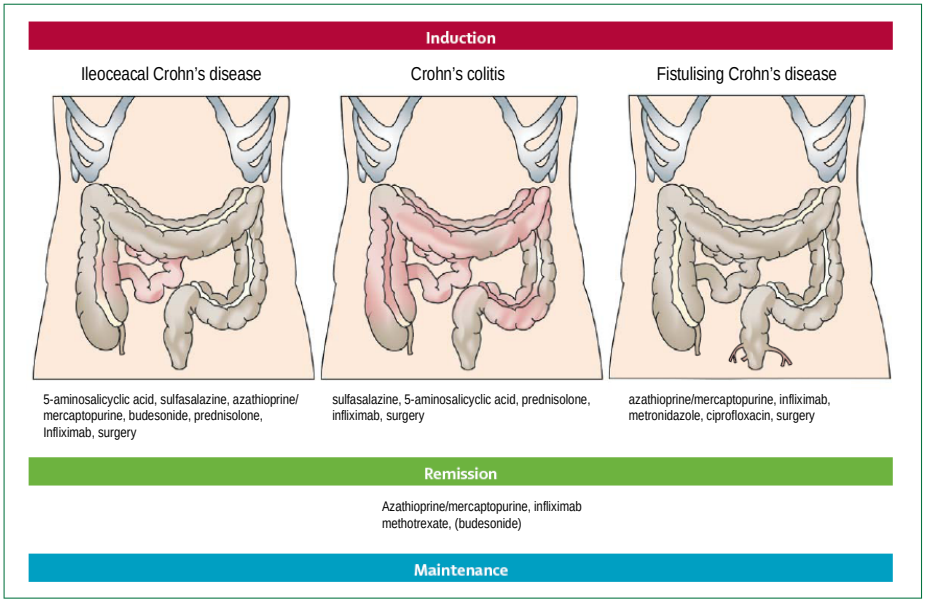


Figure 2. Management of Crohn's disease (source: Baumgart *et al.* [36])

2.5 IBD genetics

2.5.1 Family aggregation

A positive family history of IBD, primarily CD, is still the largest independent risk factor in contracting the disease. The contribution of genetic factors to IBD pathogenesis is mainly demonstrated by studies based on twins, with monozygotic and dizygotic twins showing 38.2% ± 0.2 and 4.1% ± 4.3 disease concordance* for CD respectively, as opposed to 12.1% ± 0.1 and 2.4% ± 2.3 in UC patients [37-39] (Table 3). Family aggregation for siblings (with a relative risk range of 13%–36% for CD, and 7%–17% for UC) and first-degree relatives (5%–10%) has also been reported [40-42]. In addition, the age of the diagnosis and disease phenotype can also be inherited [43]. However, the sort of IBD can be independent of the disease of the first-degree relative, indicating that both diseases can share common susceptibility loci.

Table 3. Concordance* of IBD between twins.

IBD	Monozygotic twins		Dizygotic twins		Reference
CD	8/18	38.2% ± 0.2	1/26	4.1% ± 4,.3	1
	5/25		4/46		2
	5/10		0/27		3
UC	1/16	12.1% ± 0.1	0/20	2.4% ± 2.3	1
	6/38		1/34		2
	3/21		2/44		3

1: Tysk *et al.*, 1988 [39]. N (total twins)= 25000. N_{IBD} (at least one with IBD)= 80
2: Thompson *et al.*, 1996 [38]. N= 16000. N_{IBD} = 150
3: Orholm *et al.*, 2000 [41]. N= 29421. N_{IBD} = 103

2.5.2 Susceptibility loci

IBDs are considered complex multigenic diseases in which the interaction of multiple genetic and non-genetic factors is involved. Genome-wide studies have contributed to the detection of several associated susceptibility loci, although some of the studies need replicating to confirm this [27, 44]. These genes are related to three main pathophysiological mechanisms [45]: i) an altered innate immune response to enteric microbiota or pathogens; ii) an epithelial barrier with augmented permeability; and iii) an inappropriate regulation of the adaptive immune system. A summary of the most important susceptibility genes for IBD is listed in Table 4.

Some risk loci, such as *IL23R*, *IL12B*, *HLA*, *NK2-3* and *MST1*, are common to CD and UC, whereas the autophagy* genes *ATG16L1* and *IRGM*, in conjunction with *CARD15* (formerly known as *NOD2*) are specific to CD. Fisher *et al.* [46] suggest that colonic forms of IBD (ulcerative colitis and Crohn's colitis) share pathogenic mechanisms that are distinct from small bowel inflammation (Crohn's ileitis). In turn, the association of *CARD15*, *ATG16L1* and *IRGM* with CD but not UC indicates that alterations in the intracellular processing of bacteria constitute a central feature of the pathogenesis of CD [47].

The first gene to be implicated in CD was *CARD15* [48, 49]. It encodes for an intracellular pattern recognition receptor which contains a leucine-rich repeat region that recognises bacterial muramyl-dipeptide (MDP). The recognition triggers inflammation via a nuclear transcription factor-kappaB (NF-kB) pathway. Three independent mutations represent 82% of *CARD15*-CD variants (Arg702Trp, Gly908Arg and 3020insC (1007fs)). At least one of these mutations is present in 10 –20% of Caucasian CD patients. The risk of disease development is increased 20- to 40-fold for homozygous and 2- to 4-fold for heterozygous patients. Those variant alleles have been associated with onset of the disease at a younger age, and the ileal and stricturing disease behaviour of CD patients. The severity of the disease and extraintestinal manifestations were not different for any of the *CARD15* genotypes [50, 51].

Autophagy*-related genes (*ATG16L1* and *IRGM*) have been more recently implicated in CD pathogenesis [52-54]. Autophagy* is essential for cellular homeostasis because it provides a mechanism of response to several stresses and plays an important role in the degradation of harmful exogenous or endogenous components, such as intracellular bacteria and toxins. A failure in the clearance of invasive bacteria due to alterations in autophagy* or phagosomal functions would lead to chronic inflammation.

Table 4. Most important susceptibility genes for IBD. Extended and modified from Goyette et al. [45]. ● Replicated studies.

Gene	Locus	Function	Expression	Variant(s) (genomic region)	Putative effect of the variant	Disease exclusivity	References
Innate immunity							
● *CARD15* (*NOD2*)	16q12 (IBD1)	Intracellular receptor for MDP; activates NF-kB in response to ligand	IEC, PC (small intestine), Ma, DC, Mo	3020insC (1007fs), Arg702Trp, Gly908Arg	Alter the structure of either the LRR domain of the protein or the adjacent region; decreased NF-kB activation	CD, particularly l-CD	[48, 49, 55]
CARD4 (*NOD1*)	7p14.3	Intracellular receptor for iE-DAP (a PGN-like molecule); activates NF-kB in response to ligand	Large and small intestines	ND1+ 32656*1 (rs6958571)	Affects the binding of an unknown nuclear factor	UC/CD	[56-58]
● *PPAR-γ*	3p25	Peroxisome proliferative-activated receptor γ; Nuclear receptor that is activated by aminosalicylic acid inhibiting NF-kB activity; more expressed in adipose tissue and colon and less in ileum	IEC, Ma	PPARγ1, PPARγ2, and PPARγ3. Decreased expression in UC	Impaired inhibition of pro-inflammatory cytokines, chemokines, adhesion molecules, and cell proliferation Upregulation of inflammatory mediators	CD UC	[59, 60] [61]
TLR4	9q32-q33	Transmembrane PRR, activate APCs to support Th1 cell differentiation, induction of inflammation and establishment of adaptive immunity	IEC, Ma	Asp299Gly	Modulates severity of inflammation in experimental models of colitis. Protective against CD, reduced production of bacterial flagellin specific antibody	IBD	[50, 62-67]
NF-kB1	4q23-q24	Downstream effector of signalling through the NOD and TLR receptor families	B cells, T cells and many other cell types	294delATTG	Reduced binding potential to colonic nuclear extracts and reduced activity of the mutated promoter	CD	[68, 69]
CD14	5q31.1	Co-receptor in the presentation of LPS to TLR4	Mo	2159(T/C)	Promoter polymorphism affecting interaction with CARD15 ?	IBD	[65, 70, 71]
Barrier integrity							
● *OCTN1* (*SLC22A4*) ● *OCTN2* (*SLC22A5*)	5q31 (IBD5)	Na-independent organic cation transporter 1 and Na-dependent high affinity OCT 2	IEC, Ma, T cells	L503F –207G/C	Impaired elimination of endogenous small organic cations, drugs and environmental toxins by a decrease in transporter activity and altered promoter activity	CD	[72, 73]
MYO9B	19p (IBD6)	Myosin 9B which contain a Rho-GTPase activation domain; Filament remodelling and regulation of tight junction assembly	–	Ala1011Ser (rs1545620)	Regulation of epithelial permeability	UC / (CD)	[74]
● *MDR1* (p-glycoprotein 170)	7q	(ATP)-dependent efflux pump; Elimination of xenobiotics such as bacterial products	Epithelial surface	C3435T	Altered pharmacokinetics of a wide range of drugs, including corticosteroids and immunosuppressives. Alterations in MDR1 expression and activity may result in altered intestinal absorption	Refractory cases of IBD	[75, 76]
● *DLG5*	10q23	Encodes for an scaffolding protein that helps to maintain epithelial integrity	–	G113A	Probably impedes scaffolding of DLG5	CD/ (UC)	[77]

Table 4. Most important susceptibility genes for IBD. Extended and modified from Goyette *et al.* [45] (*Continued*)

● *PTGER4*	5p13.1	Prostaglandin receptor EP4; regulates epithelial barrier function	IEC	(rs4495224)	Increased PTGER4 expression	CD	[78]
Adaptive immunity							
TNF-α	6q21-p22	Proinflammatory cytokine		–857C/T –308G/A	May affect the transcription and expression of TNF, and further influence its biological function.	CD stenosing and penetrating UC	[79-81]
● *IL23R*	1p31	Encodes a crucial subunit for the *IL23* receptor	Activated DC, phagocytic cells	(rs1004819) (rs2201841) (rs10889677) (rs11209026) (rs7517847) (rs10489629) (rs1343151)	Alters the efficiency of *IL23R* signalling; the additional disease-association signals in the *IL23R* region may result from polymorphisms that alter *IL23R* expression	enhances susceptibility IBD protective SNPs for IBD	[82]
MHL locus (*HLA* class I and II genes)	6p21 (IBD3)	Antigen presentation	Leukocytes	HLA-DRB1*0103 HLA-DRB1*1502 HLA-DRB1*07, HLA-DRB1*0103, HLA-DRB1*04, and HLA-DRB3*0301	-	UC and colonic CD CD	[83, 84]
● *ATG16L1*	2q37.1	Encodes autophagy*- related 16-like protein 1	IEC, APCs, CD4/8 T cells, B1 cells and B memory cells; colon and small bowel	(rs2241880) T300A	Altered phagosome function affecting handling of commensal flora?	CD, especially I-CD	[52, 85]
● *IRGM*	5q33.1	Encodes immunity-related GTPase family M; related to autophagy*; required for mycobacterial immunity	-	Altered expression, not single mutation	Alteration in *IRGM* regulation that affects the efficacy of autophagy*	CD	[53, 54]
NCF4 (p40^{Phox})	4p13	NADPH oxidase activity and generation of reactive oxygen species upon phagocytosis	Hematopoietic cells (Ne, Mo, Eo, MC and basophils)	(rs4821544)	Altered phagosome function affecting handling of commensal flora	IBD	[85]

Abbreviations: DC, dendritic cells; Eo, eosinophils; IEC, intestinal epithelial cells; Ma, macrophages; MC, mast cells; Mo, monocytes; Ne, neutrophyles; PC, Paneth cells; MDP, muramyl-dipeptide; PGN, peptidoglycan; LPS, lipopolysaccharide; PRR, pattern recognition receptor; APC, antigen presenting cells; NOD, nucleotide-binding-oligomerisation domain; TLR, toll-like receptor; iE-DAP, g-D-glutamyl-meso-diaminopimelic acid.

3 Microbial-host interactions in the intestinal mucosa

3.1 Composition and metabolic activity in an enteric microbial community

Despite the clinical importance of microbiota, little is known about the composition, activity and function of human microbiota. Three main factors hinder an accurate description of such a complex environment: i) the high degree of complexity of the microbial community in the intestinal tract, ii) the difficulty in obtaining samples from along the entire GI tract from healthy individuals, and iii) the limitations of available techniques.

On top of the 400 species found inhabiting the GI tract by culturing, more than 1000 phylotypes have been obtained by various molecular approaches to date (for review, [86]). In 1998, Zoetendal *et al.* [87] first described the intestinal microbiota using PCR-DGGE, and was followed by Suau *et al.* (1999) [88], who used a cloning-sequencing approach. These two first studies and most of the ones that followed were derived from fecal samples, given how easy it is to sample large numbers of individuals. Some years later, the bacterial composition and abundance in feces was reported to differ from biopsies [89, 90]. Studies using culture-independent methods and tissue samples were then carried out, first by Hold *et al.* (2002) [91], and later, but more extensively, by Eckburg *et al.* (2005) [92]. More than 15,000 sequences obtained from the human GI tract are in public databases, of which 13,335 originate from Eckburg *et al.*, 2005 [92]. Culture-independent methods are providing more data about the ecological features of the human microbiome*, although a combination of techniques is still necessary to achieve a better picture of the community. For a review that integrates results from both culture-based and molecular-based studies in the period 1998-2006, see reference [86].

It has been said that our intestinal tract is colonised with up to 10^{14} microorganisms. This population exceeds in size any other microbial community associated with the body and is 10 times greater than the total number of human cells. The varying conditions along the gut determine the abundance of indigenous microorganisms in each zone. Bacterial concentrations in the oesophagus range from 10^2 to 10^3 cfu cm^{-2} of tissue. In the stomach, numbers decrease to values that range from 10^1 to 10^2 cfu ml^{-1} of aspirate, followed by an increase along two-thirds of the small intestine (from 10^2 to 10^4 cfu ml^{-1} of

aspirate). Lactobacilli and streptococci are the main colonisers of these zones of the GI tract. In the distal ileum, bacterial concentration increases to 10^7-10^8 cfu g^{-1}, reaching its highest abundances in the colon (10^{11}-10^{12}) ([93] and references therein). Although this data is extensively used and accepted by many scientists, it is based on culture-based methods. Different counts have been obtained using molecular methods, with higher bacterial concentrations being found in terminal ileum mucosa than in colonic mucosa (1.3×10^{12} 16S rRNA gene copies per gram of mucosal tissue in the terminal ileum against 2.5×10^{10} in the colon) [94].

Although the gut is a highly bacteria-colonised environment, the diversity of phyla that compose it is low. The vast majority of phylotypes fall into only two out of the 70 existing bacterial divisions. Those divisions are the phyla *Firmicutes* – mainly *Clostridium coccoides* (cluster XIVa) and *Clostridium leptum* (cluster IV) – and *Bacteroides*, which account for 57 - 82% and 16 - 31% of the phylotypes of several cloning-sequencing based studies respectively [88, 91, 92, 95, 96]. *Proteobacteria* (<10%), *Actinobacteria* (<5%), *Fusobacteria* (~1%), *Verrucomicrobia* (~0.2%), *Spirochaetes* (two phylotypes found), *Lentisphaerae* (one phylotype) and *Cyanobacteria* (one phylotype) are additional phyla that have been detected, either by culture-based or culture-independent methods [86]. Members of Eukarya (nine phylotypes belonging to *Ascomycota*), Archaea (*Methanobrevibacter smithii*, *Methanobrevibacter ruminantium* and *Methanosphaera stadtmanae*) and more than 1200 viral genotypes have also been detected within this complex community. Figure 3 shows the phylogenetic relationship of the 1151 distinct intestinal phylotypes used in the meta-analysis* performed by Rajilić-Stojanović *et al.* [86], who combined non-redundant sequences retrieved from the GenBank from both culture-based and culture-independent methods.

Many studies on bacterial composition that use molecular profiling techniques, such as TGGE and DGGE (temperature/denaturing gradient gel electrophoresis), have provided evidence that:

- Although there are some species in common among individuals, such as *Faecalibacterium prausnitzii*, the entire intestinal microbiota is host-specific [87, 90, 92, 97-99].
- The intestinal microbiota of adults is stable over time, at least for two years [87, 99, 100].

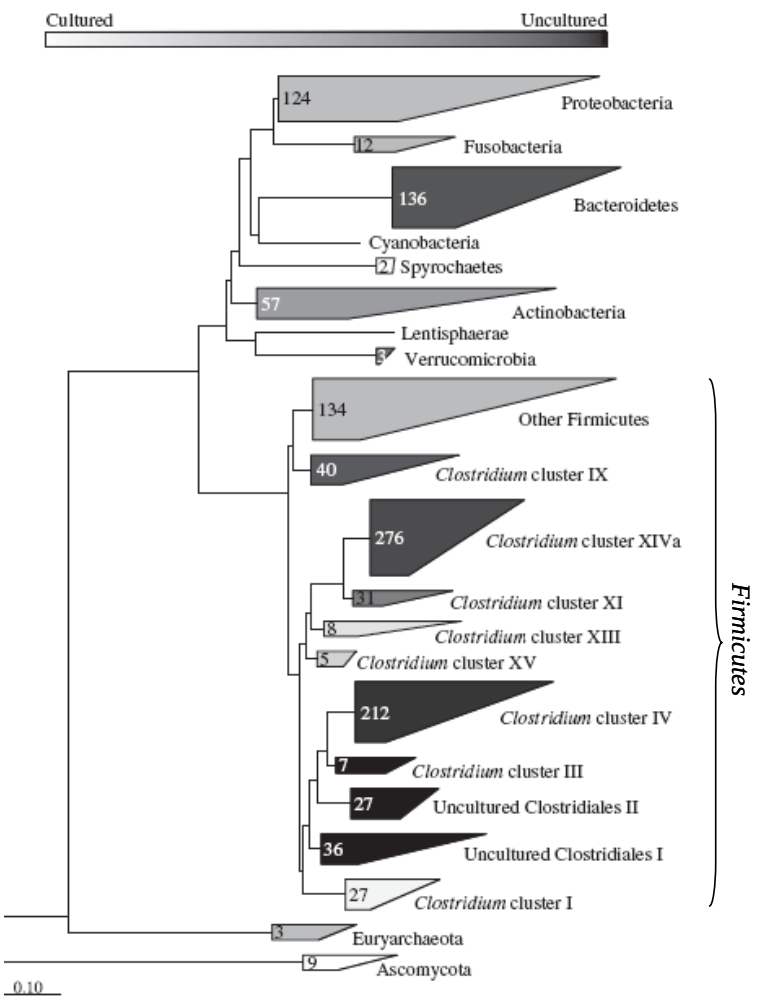


Figure 3. SSU rRNA-based phylogenetic tree of 1151 intestinal selected phylotypes by Rajilić-Stojanović *et al.* [86]. Phylotypes obtained by cultivation-based works (white filling) and molecular techniques (black filling) are integrated and the number of distinct phylotypes among each phylogroup are given.

- In newborns, there is a succession of microorganisms during the first days of life. Intestinal mucosa is first colonised by facultative anaerobes, while strict anaerobes gradually become more abundant and diverse [101].
- Fecal microbiota differs from tissue-associated microbiota [90, 92].
- Microbial composition is similar along the distal ileum, colon and rectum [90, 92, 98], but differs from that colonising the upper GI tract [96]. Essentially, aerobes and facultative anaerobes, such as *Streptococcus, Lactobacillus* and *Enterobacteriaceae*, colonise the duodenum and jejunum, whereas *Firmicutes* and *Bacteroidetes* are more abundant in the distal ileum, colon and rectum.

Many factors shape microbial diversity in the human intestinal gut [102]. For example, the chemical niches available such as pH or redox, physical factors such as intestinal peristalsis, and other factors such as microbial competition, host pressure (mutualistic microorganisms are favoured), and the host's immune system. Studies using animal models have contributed to the knowledge of exogenous and endogenous factors contributing to host-microbe specificity. In particular, Rawls *et al.* [103] elegantly demonstrated that vertebrates can

modulate their own microbiota. They exchanged the microbiota of zebra fish and mice and confirmed that the host recipient shaped the composition of the non-native microbiota to one closer to its native community, which is evidence of strong host pressure.

Despite host-microbe specificity, the overall metabolic activity of intestinal microbiota is similar amongst subjects. Using a metagenomic approach, Gill *et al.* [104] demonstrated that the human gut microbiome* is enriched in genes involved in starch and sucrose metabolism, plus the metabolism of glucose, galactose, fructose, arabinose, mannose, and xylose. More recently, Chassard *et al.* [105] have determined the population levels of the main metabolic groups. They observed that starch-degrading bacteria represented 10.1% of total viable counts, while mucin degraders accounted for 5.1%, and proteolytic bacteria for 1%. Among the fibre-degrading population, xylanolytic bacteria accounted for 2.6%, and cellulolytic bacteria for 0.16% of total bacteria. The fermentation of many compounds requires the cooperation of different metabolic groups linked in a trophic chain. For example, glycans present in the mucus are degraded by primary fermenters into short-chain fatty acids (SCFAs) such as acetate, propionate and butyrate, plus gases such as H_2 and CO_2. The majority of SCFAs are absorbed by host cells. SCFAs like butyrate are the principal energy source for colonocytes and may fortify the intestinal barrier by stimulating their growth. Butyrate-producing bacteria (mainly composed of *Faecalibacterium prausnitzii*, *Roseburia* spp. and *Eubacterium* spp.) are therefore important for intestinal health. Hydrogen can be removed from the intestine in different ways: by methanogenesis, sulphate reduction, and homoacetogenesis. Methanogenic *Archaea* are more abundant in those subjects that are methane excretors (30-40% of the human population), and reductive acetogenic bacteria in non-methane excretors. Sulphate reducing bacteria are present in similar abundance in both groups of subjects [105]. Genes involved in the synthesis of amino acids and vitamins, as well as those involved in the detoxification of xenobiotics, have also been detected in the human microbiome*.

3.2 The homeostatic situation in a healthy gut

Host-microbe interactions occur principally along the mucosal surfaces [106]. Mucosal surfaces are interfaces that separate the immune system of the external microbial world physically and chemically. Spread along the intestinal mucosa, gut associated lymphoid tissue (GALT) represents the largest part of the body's immune system (about 70%), in the small intestine forming Peyer's patches, and

in the colon lymphoid follicles (Figure 4). Moreover, the largest amounts of mucosa-associated bacteria reside in the human gut (approximately 10^{14} gastrointestinal bacteria versus a total of 10^{13} cells in the body [93]). Such a situation must be very delicately regulated to avoid a state of constantly uncontrolled inflammation [107-109]. The immune system must be able to be tolerant of commensal bacteria, but at the same time ready for an active response to pathogens. In normal hosts, commensal bacteria activate sequenced homeostatic responses by epithelial cells, macrophages, dendritic cells, T lymphocytes and B cells. During microbial colonisation, the immune system matures, and the host achieves a tolerance to commensal bacteria. In addition to postnatal intestinal maturation, intestinal microbiota play other important roles in the physiology of the human gut, including functions such as nutrient absorption, mucosal barrier fortification, xenobiotic metabolism, and angiogenesis [108].

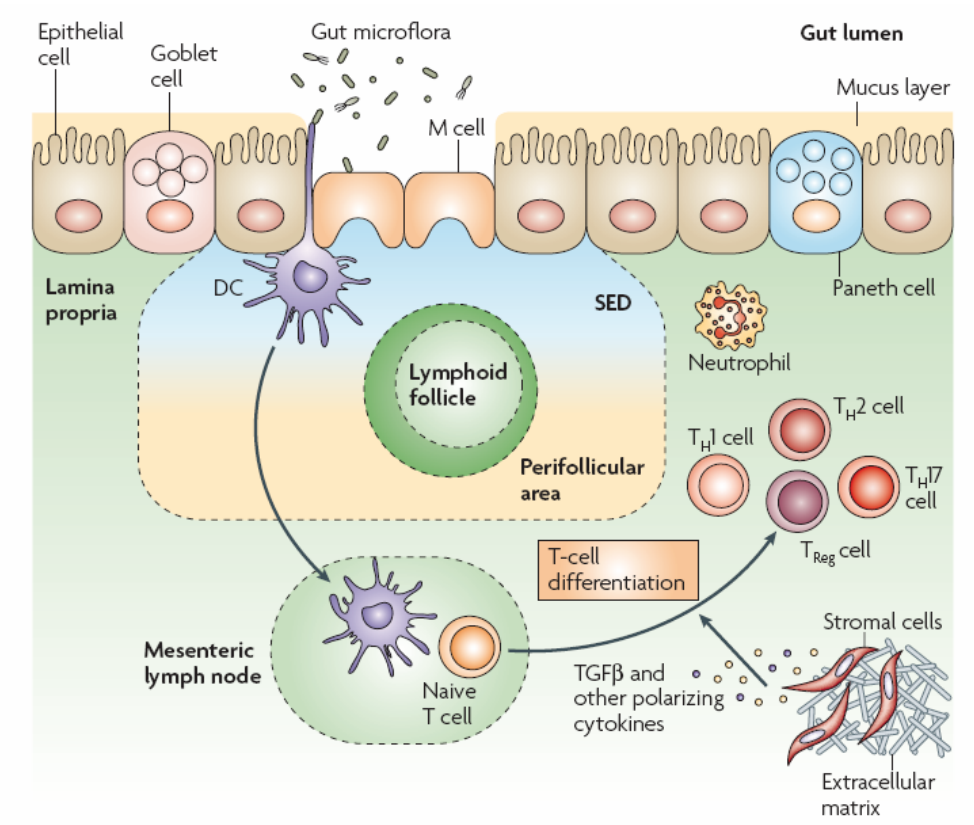


Figure 4. Schematic view of the main features of the intestinal immune system. From Cho, 2008 [47]. The epithelial surface is comprised of intestinal epithelial cells (absorption and secretion), goblet cells (formation of mucus) and Paneth cells (secretion of antimicrobial defensins). Microfold cells (M cells) and dendritic cells (DCs) sample intestinal luminal contents. In the presence of pathogens, activated DCs migrate to the mesenteric lymph nodes to further activate naïve T cells, which then undergo differentiation under the influence of factors released by DCs and other stromal elements. Abbreviations: SED, subepithelial dome; TGFβ, transforming growth factor-β; T_H, T helper; T_{Reg}, T regulatory.

Homeostasis is made feasible by a down-regulation of bacterial receptors and their ligands, as well as a low ratio of pro-inflammatory/anti-inflammatory cytokines and the stimulation of protective molecules that mediate mucosal barrier function [110]. Specialised regulatory T cells take part in this process of tolerance through the functional inactivation of immune reactive cells (clonal anergy). Smythies *et al.* [111] have described the macrophage population in a homeostatic situation. In normal hosts, these cells are phagocytic but unresponsive in an inflammatory way. They are unable to initiate inflammatory responses to bacteria because they are in a state of inflammatory anergy. In the absence of inflammation, macrophages show a reduced expression of receptors for antigens and a down-regulation of cytokine production, despite the proximity of immunostimulatory antigens, contributing to the oral tolerance.

3.2.1 The influence of commensal microbiota

There is a mutualistic relationship between resident microbiota and the body [109, 112]. Probably the most important beneficial contributions of commensal bacteria to the host are i) a defence against pathogens [106], ii) the promotion of postnatal gut and immunity development [108], and iii) the digestion of dietary products [113]. Germ-free and recolonised animal models have been very useful in gaining an understanding of some of the aspects of the interaction between bacteria and the host. Defence against infection can be mediated directly by the bacterial community via displacing pathogens through competition for nutrients and epithelial binding sites, as well as indirectly by promoting the activation of host defences as, for example, by producing antimicrobial peptides. Some of the ways in which resident microbiota can promote human health are listed below (adapted from Sears, 2005 [114]).

- Polysaccharide utilisation and nutrient release
- Enhanced fat storage, production of short-chain fatty acids
- Induction of mucosal glucose transporters
- Induction of villous capillary formation
- Breakdown of carcinogens and synthesis of biotin, folate and K vitamins.
- Induction of proteins of innate immunity; i.e. Angiogenin-4
- Contribution to mucosal homeostasis and repair capacity
- Stimulation of secretory immunoglobulin-A (IgA) production
- Induction of gut-associated lymphoid tissue development
- Promotion of diversification of lymphoid populations and immunoglobulin genes

3.2.2 The epithelial surface

Intestinal epithelial cells provide several lines of defence against bacterial antigens. First, they constitute an efficient physical barrier that impedes penetration of macromolecules and intact bacteria [27]. Pre-epithelial molecules and tight junctions contribute to this physical isolation. Second, epithelial cells are the primary site of antigen recognition due to their expression of extracellular and intracellular receptors, mainly toll-like receptors (TLR) and nod-like receptors (NLR), which are the nodal point between bacteria and the dense network of innate and adaptive immune cells. The ligation of a wide range of microbial adjuvants and antigens activates two main signal transduction pathways, nuclear transcription factor-kappaB (NF-κB) and mitogen-activated protein kinase (MAPK), which both lead to the expression of a number of pro-inflammatory and anti-inflammatory genes (Figure 5).

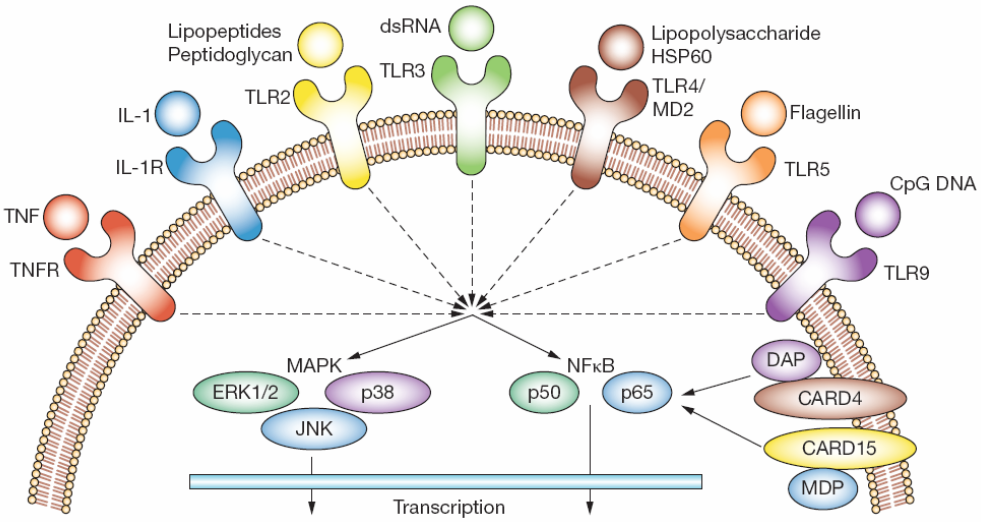


Figure 5. Recognition of microbial adjuvants by extracellular (TLRs) and intracellular (CARDs) pattern recognition receptors. The ligation activates two main signal transduction pathways, NF-κB and MAPK. Abbreviations: CARD, caspase recruitment domain family member; CpG DNA, DNA containing cytosine–guanine repeats linked by phosphodiester bonds; DAP, diaminopimelic acid; dsRNA, double-stranded RNA; ERK, extracellular signal regulated kinase; HSP60, heat shock protein 60; IL-1, interleukin-1; IL-1R, interleukin 1 receptor; JNK, c-Jun amino-terminal kinase; MAPK, mitogen-activated protein kinase; MDP, muramyl dipeptide; NFκB, nuclear factor κB; P38, mitogen-activated protein kinase 1; P50, subunit of NFκB that forms a heterodimer with P65; P65, NFκB subunit; TLR, Toll-like receptor; TLR4/MD2, complex of Toll-like receptor 4 and MD2, a molecule that confers responsiveness on lipopolysaccharide; TNF, tumour necrosis factor; TNFR, TNF receptor. From Sartor R, 2006 [28].

In addition to columnar epithelial cells, other specialised cells that are dispersed along the crypts* play a role in protecting against microorganisms and the promotion of mucosal repair. **Paneth cells** are situated in the base of the crypts*, mainly in the terminal ileum. These cells produce a battery of antimicrobial

peptides to regulate the density of microorganisms. These proteins are wide acting and effective against gram-positive and gram-negative bacteria, fungi and viruses. Defensins*, lysozyme and phospholipase A2 are some of these antimicrobial proteins. Defensins can be divided into two families: α-defensins, which occur only in the small intestine, and β-defensins, which are expressed on all mucosal surfaces of the GI tract. **Goblet cells** are granular simple columnar epithelial cells whose main function is to synthesise mucin glycoproteins that confer a gel-like layer on the epithelial surface. Their physiology has been shown to play a protective role against colitis by preventing the firm adhesion of bacteria and distancing the microbiota from the epithelial surface [115]. Microfold cells, also known as **M cells,** are epithelial cells that cover the follicle-associated epithelium of Peyer's patches. They lack microvilli and the mucus layer is thinner or absent on their apical surface. They are specialised in sampling antigens (macromolecules and microorganisms) from the gut lumen and delivering them via transcytosis to lymphoid tissue cells situated in a pocket-like zone located on their basolateral side. Some pathogens exploit M cells as a route of invasion.

3.2.3 Antigen recognition

Luminal antigens are sampled by professional and non-professional **antigen-presenting cells** (APCs) via extracellular and intracellular receptors, called **pattern recognition receptors** (PRR). Intestinal enterocytes and M cells are non professional antigen-presenting cells. Their dialogue with immunologic cells is mediated by the major histocompatibility complex (MHC) proteins, which are not constitutively expressed. MHC proteins are expressed only upon stimulation of the non-professional APCs by certain cytokines. Professional antigen-presenting cells express the whole arsenal of PRRs and co-stimulatory molecules. **Dendritic cells** are the key APCs; they expose their dendrites to the intestinal lumen to sample bacterial components. Two types of PRRs are involved in antigen recognition in the GI tract: membrane associated **toll-like receptors** (TLR) and cytosolic **nucleotide-binding-oligomerisation-domain** (NOD) proteins. Several cell types along the GI tract present these PRRs, which participate in bacterial recognition in at least four ways: i) recognition of pathogen associated molecular patterns (PAMPs); ii) connection with the GI lumen; iii) linkage with the innate and adaptive immune system by the induction of pro/anti-inflammatory molecule secretion; and iv) induction of other antimicrobial effector pathways [116].

Mammalian TLRs are transmembrane molecules composed of three common structural domains (Figure 6). In the extracellular zone, a divergent ligand-binding domain composed of multiple leucine rich repeats constitutes the

recognition zone. Next, a short transmembrane region connects the recognition zone to the highly homologous cytoplasmic Toll-interleukin 1 receptor (TIR) domain, which is essential for the initiation of downstream signalling cascades [117]. Many cell types of the intestinal mucosa (mature and immature epithelial cells), the lamina propria (monocytes, macrophages, and dendritic cells) and the submucosa (myofibroblasts, endothelial cells, and adipocytes) express, inducibly or constitutively, several TLRs [116]. Although these receptors are specific to certain ligands, the presence of a variety of them (at least 11 known to date) allows the recognition of most PAMPs (Figure 7).

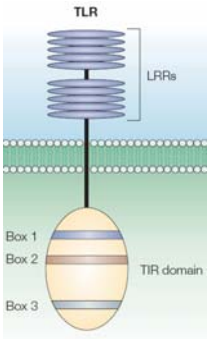


Figure 6. Toll-like receptor structure. From Akira and Takaeda, 2004 [118]. Cytoplasmic toll-interleukin 1 (TIR) domain is highly conserved. TIR contains three regions (boxes 1, 2 and 3) of high homology with interleukine-1 receptor. Extracellular toll-like receptor (TLR) domain contains tandem repeats of leucine (LRRs, leucine rich repeats) and is variable, which allows the recognition of different antigens.

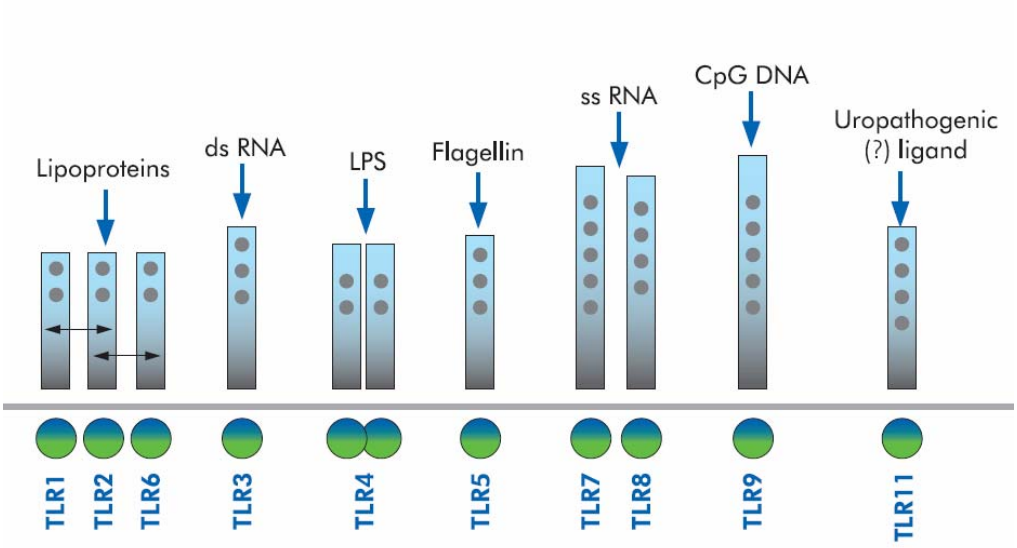


Figure 7. Main pathogen associated molecular patterns (PAMPs) recognised by different TLRs. From Cario, 2005 [116].

More than 20 different NOD-like receptor (NLR) proteins are known to date, with a common structure comprising three domains: i) a carboxy terminal domain consisting in leucine-rich repeats, ii) a central nucleotide binding and oligomerisation domain, and iii) a variable amino terminal effector binding domain (Figure 8). Depending on the type of effector binding domain, NLRs are divided into subfamilies ([119] and references herein), as for example the NOD

subfamily, whose effector binding domain corresponds to a caspase activation and recruitment domain (CARD). In response to distinct bacterial ligands, both NOD1-CARD4 and NOD2-CARD15 have the major responsibility for the intestinal regulation of proinflammatory signalling through NF-κB. In particular, NOD2 recognises peptidoglycan (PGN) and muramyl dipeptide (MDP), whereas NOD1 senses a Gram negative PGN derivative [120, 121]. In the intestine, the expression of NOD1 and NOD2 is low but increases in inflammatory situations [122, 123]. The highest expression of NOD2 is found in Paneth cells [124].

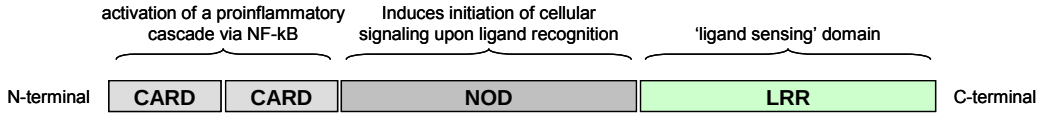


Figure 8. Diagram of NOD2 receptor structure. Abbreviations: CARD, caspase activation and recruitment domain. NOD; central nucleotide-binding and oligomerisation domain. LRR: leucine-rich repeats.

3.2.4 Innate Immunity and Adaptive Immunity

Innate immunity is an inborn system that operates as a first line of defence against microorganisms. It comprises circulating cells such as neutrophils and monocytes, as well as resident intestinal immune cells (dendritic cells, Paneth cells and myofibroblasts) [109, 125, 126]. The intestinal epithelium, including the antimicrobial peptides synthesised in some zones (defensins* and cathelicidins), are also included as elements in the innate immunity. In contrast, adaptive immunity generates a slow and more targeted response, basically by means of B and T lymphocytes, involving antigen-specific recognition and immune memory. MHC proteins of antigen presenting cells are the nexus for the adaptive immunologic response.

Dendritic cells have primary responsibility for controlling both the innate and the adaptive immune responses in the intestinal mucosa. They regulate the activation of natural killer cells, induce the production of IgA^+ B cells, and mediate a balanced differentiation of immature T cells (Th0) into effector (Th1, Th2, Th17) and regulatory (Th3) T cells. **Intestinal macrophages** are strategically situated in the lamina propria, under the epithelial monolayer, from where they can survey the possible crossing of bacteria. Intestinal macrophages are different to those from the systemic immune system. They are active to phagocyte but show low

levels of inflammatory responses, probably due to a low expression of PPRs. This adaptation helps them to be tolerant to commensal bacteria. **Myofibroblasts**, have recently been proposed as possible innate immune mediators because they are able to recognise microbial antigens, as well as to respond by producing proinflammatory mediators after antigen-receptor recognition [126].

In situations of homeostasis, innate immunity is constantly under a state of "controlled activity", synthesising antimicrobial peptides in order to control the development of the intestinal microbiota, and sampling their antigens. Secondary effector cells, such as granulocytes, mast cells, natural killer cells, and macrophages, remain in their compartments unless there is an infection or a deregulated situation [44].

3.3 Disruption of homeostasis in Crohn's disease (*Host factors*)

3.3.1 Epithelial barrier deficiencies in Crohn's disease

During the '90s, the role of increased **epithelial permeability** across the gut (leaky gut) gained support as a factor involved in IBD pathogenesis ([45] and references therein). It has been shown that there is an abnormal intestinal permeability in the mucosa of IBD patients and their first-degree relatives [127, 128]. The increased permeability precedes the onset of the disease and confers a family risk [129]. In addition, changes in tight junctions of enterocytes, by a down-regulation of junctional complexes, have been observed in IBD biopsies [128, 130]. Recently, Buhner *et al.* [131] have proposed that *CARD15* 3020insC (1007fs) mutation could be one genetic factor involved in the impairment of intestinal barrier function. However, the absence of this mutation in 100% of CD patients and its presence in some healthy relatives indicates that additional genetic or environmental factors are necessary to develop the disease. Another genomic region which it has been suggested plays a role in epithelial barrier function is located in the susceptibility locus 5p13.1 and regulates the expression of the prostaglandin receptor EP4 (PTGER4)[78]. PTGER4 works at different levels: maintaining mucosal integrity, suppressing innate immunity, and down-regulating the proliferation and activation of CD4+ T cells. Finally, many other cells and cellular components such as T cells, neurons, tumour necrosis factor, and interleukin-3 might mediate intestinal permeability [132-135].

Defects in mucus production have also been detected in IBD patients [136, 137], possibly due to an altered expression of goblet cells [138, 139]. However, it is thought that mucus play a protective rather than an etiological role in IBD.

3.3.2 Disturbed innate immune mechanisms in bacterial recognition and clearance

The role of innate immunity in both the early events and the perpetuation of CD has come to be thought of as more important over the last few years, whereas defects in adaptive immunity are now thought to be a secondary event [140].

There is evidence that tolerance to commensal bacteria is broken under intestinal inflammation. Duchmann *et al.* [141] observed that cells obtained from the gut of IBD patients exhibit inflammatory responses when cultured with both resident and heterologous bacteria, whereas cells from controls only react to the presence of heterologous bacteria. The mechanisms involved in this pathogenic behaviour are still unclear. Nonetheless, more discoveries are being made in this field.

First, there is an **altered expression of PPRs** in the enterocytes and professional APCs of IBD patients. The receptor for dsDNA, TLR3, is down-regulated in the epithelial cells of CD patients, whereas TLR2 and TRL4, whose PAMPs are lipopeptides, peptidoglycans, lipopolysaccharides, and HSP60, are overexpressed in active IBD throughout the lower GI tract [142]. In addition, myeloid dendritic cells exhibit a polymorphism associated with IBD in the TLR4 receptor (*TLR4*-Asp299Gly) which could trigger deficiencies in antigen recognition [64]. However, Lakatos *et al.* [50] did not find such any association between *TLR4*-Asp299Gly and CD in their study of three Hungarian cohorts. An overexpression of TLR2 and TLR4 in the dendritic cells of CD and UC patients has also been reported, as well as high levels of the CD40 maturation/activation marker IL-12, and IL-6 in the inflamed tissue of CD patients [143]. In IBD, dendritic cells are activated, expressing more PRRs and producing more proinflammatory cytokines. However, the pathogenic relevance of TLR dysregulation is not directly associable to IBD since such an imbalance can be either a cause or a consequence of the immune imbalance that characterises the inflammatory process.

Second, Wehkamp *et al.* [144] have recently reported a **reduction in α-defensin production**, mainly HD5 and HD6, by Paneth cells in intestinal mucosal extracts of CD patients. The decrease in α-defensins was specifically detected in CD patients with ileal affectation and regardless of the degree of inflammation. Although there is not a clear explanation for these reduced levels, a host factor seems to be the most feasible cause. Consistent with this, the reduced expression of α–defensins correlates with *CARD15* 3020insC (1007fs) mutation [145, 146]. Nevertheless, only a third of patients with defensin deficiency had the *CARD15* polymorphism, indicating that alternative genes may be involved in the regulation

of α–defensin expression. Recently, the same group of investigators has found a new link between IBD and defensin* deficiency, the Wnt signalling pathway transcription factor Tcf-4, which is independent of *CARD15* [147]. Tcf-4 is a known regulator of Paneth cell differentiation. Wehkamp *et al.* observed a reduced expression of Tcf-4 in the ileal tissue of Crohn's ileitis patients in comparison with Crohn's colitis and UC patients. Moreover, other host-independent factors, such as luminal or mucosa-associated bacteria, may modulate the expression of α-defensins, adding new variables to explain the differential α-defensin release in CD patients [148, 149]. Although no differences in α-defensin secretion characterise Crohn's colitis patients, it has been observed that patients with this disease phenotype have a reduced expression of β-defensins, in particular HBD2, at the level of the colon [150, 151]. The β-defensin gene cluster is located on chromosome 8p23.1, a highly polymorphic region, which has been suggested as a susceptibility locus for IBD. A lower DNA copy number of genes coding for HBD2 in Crohn's colitis patients explains the reduction in expression of β-defensin in comparison to healthy controls, UC patients and CD patients with ileal involvement [152].

Third, an **overexpression of *CARD15*** has been linked to CD. Berrebi *et al.* [122] demonstrated that the expression of *CARD15* was higher in CD patients than in controls, and that this was not restricted to mononuclear cells but also, abnormally, to intestinal epithelial cells. Later, Ogura *et al.* [124] observed that, besides monocytes, *CARD15* is expressed in Paneth cells. It is not clear whether bacteria, further host factors or a combination of both influences this deregulation. Consistent with this and downgrading the etiological role of *CARD15* deregulation, Rosenstiel *et al*. have shown that TNF-α can stimulate intestinal epithelial cells to upregulate the *CARD15* gene [123]. In addition to differences in expression, three **polymorphisms in *CARD15*** have been linked to CD, two single nucleotide polymorphisms (SNPs) – Arg702Trp and Gly908Arg – and one frameshift mutation – 3020Cins (1007fs) [48, 49, 153]. For more details see Section 2.5.2. These variants are coded in the leucine-rich repeats domain of NOD2, which suggests that *CARD15* mutations would be involved in impaired ligand recognition and, consequently, an anomalous inflammatory response [154]. Functionally, it has been demonstrated that *CARD15* variants are deficient in their recognition of MDP [121], as well as in the activation of NF-κB [49, 124, 155]. Dendritic cells with the *CARD15* 3020insC (1007fs) variant have reduced cytokine responses after stimulation with different TLR ligands in combination with MDP or PGN [156-158]. However, these findings, pointing to a loss-of-function, lead to a contradiction in the general concept of hyper responsiveness in

CD patients exhibiting increased NF-κB activity in inflamed clinical specimens against luminal bacteria [159]. Moreover, some studies based on murine models point to a gain-of-function phenotype [159, 160], which adds to the controversy around this question. Although the *CARD15* contribution to CD pathogenesis has been extensively studied, the underlying mechanism is not completely understood. Recently, Strober *et al.* [161] have reviewed this study area (Figure 9). They give two major explanations for the paradox. On the one hand, they suggest *CARD15* mutants could have a host defence defect, for example defective α-defensine production, which would enhance microbial colonisation. This hypothesis is based on the studies previously mentioned by Wehkamp *et al.* [146] and Kobayashi *et al.* [145]. In its turn, this over-colonisation would stimulate the NF-κB pathway by *CARD15* independent mechanisms. On the other hand, they suggest that there is synergistic behaviour between the intracellular receptor CARD15 and the extracellular receptor TLR2 that mediates the immune response modulation. This hypothesis, which is not mutually exclusive with the previous, is based on the works of Watanabe *et al.* [160, 162]. PGN is present in the cellular wall of the vast majority of microorganisms, and is sensed by the TLR2 of the intestinal mucosa APCs. This ligation activates the NF-κB and hence leads to downstream pro-inflammatory cytokine production. PGN is also digested intracellularly in the endosome into MDP, which is sensed by and activates *CARD15*. In normal hosts, *CARD15* activation would initiate a mechanism of inhibition of the NF-κB pathway that would further cause a down modulation of TLR-induced inflammatory response. In this sense *CARD15* would participate in a mechanism of bacterial tolerance. In *CARD15* deficient CD patients, *CARD15* modulation would be impaired (not NF-κB inhibition) and, consequently, the TLR-induced inflammatory response would be enhanced.

Fourth, **non-professional APCs**, such as epithelial cells, can become potent effector-T-cell activators in IBD patients. Generally, epithelial cells from IBD patients acquire an activated phenotype with high levels of MHC in the presence of proinflammatory cytokines, such as IFN-γ and TNF-α [163]. In addition, IBD enterocytes can express co-stimulatory molecules, such as B7h and B7-H1, which might transform them **into functional APCs** [164].

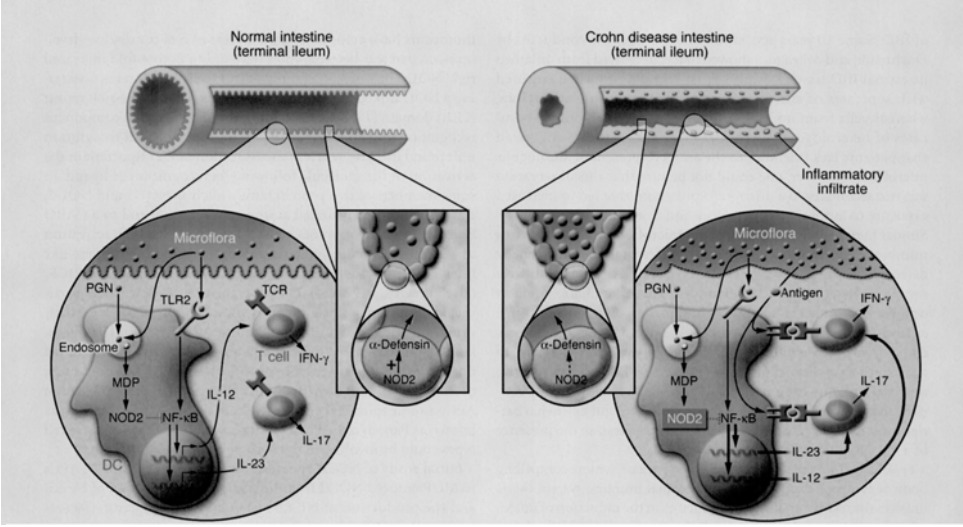


Figure 9. Hypothesis of the implication of NOD2 deficiencies in Crohn's disease pathogenesis. From Strober *et al.*, 2007 [161].

3.3.3 Unbalanced adaptive immune response

CD patients exhibit an increased number of adaptive immune cells in their intestinal mucosa. While during acute phases there is an increase in neutrophils, there is a high concentration of T lymphocytes in the CD mucosa irrespective of the state of inflammation. This suggests that mucosal immunity could play an active role in the etiopathogenesis of CD, and is probably related to the chronic nature of the disease [165]. However, whether the disease is driven by an aberrant response to normal commensal microbiota or to a specific pathogen still remains uncertain. Gewirtz proposes that T-cell unbalance in CD could be mediated by certain bacterial flagellins that would be detected by TLR5 situated on the basolateral side of intestinal epithelial cells. It has been demonstrated that when flagellin is administered systemically in mice, it serves as an adjuvant for both T-cell and antibody responses [166, 167]. This shows that flagellin can stimulate both innate and adaptive immune responses. Alternatively or simultaneously, the high concentration of activated T-cells in CD mucosa could be due to a defective clearance by resistance to apoptosis [168]. There are some biological therapies that are currently being used for IBD that block the activation cycle and induce T cells to apoptosis [169]. Not only are there quantitative but also qualitative differences in adaptive immune responses in IBD patients. IBD involves an increased effector/regulatory T-cell ratio as well as abnormal patterns of cytokine production and immune cell responses [170, 171]. In CD patients, IL-23 [172] and Tbet [173] mediate the differentiation of naive T-cells (Th0) into activated Th1 cells [174], which secrete pro-inflammatory cytokines such as IL-12, IL-18, TNF-like, and IFN-γ. In turn, IFN-γ stimulates macrophages to release IL-1, IL-6

and TNF-α. In UC patients, natural killer T-cells and lamina propria T-cells secrete high amounts of the Th2 cytokine IL-13, and to a lesser extent IL-5, which have a potentially cytotoxic effect on epithelial cells [132, 170]. More recently, some evidence indicating a role of Th17 cell proinflammatory responses in IBD has been reported. Brand *et al.* [175] observed an increase in IL-22 mRNA and IL-17F (also referred to as IL-24) in CD patients, as well as increased levels of IL-17A that correlated with IBD [176-178]. It is known that IL-23 mediates the activation of Th17 cells [178]. Interestingly, some associations have been found between variants in the IBD-susceptibility loci *IL23R* and the increase in IL-22 production [179].

4 Role of intestinal microbiota in Crohn's disease (*Microbial Factors*)

4.1 Evidence of bacterial implication in the pathogenesis of Crohn's disease

There is some evidence to implicate intestinal microbiota in the pathogenesis of CD. Clinical improvement in CD patients by diversion of the fecal stream after surgery, the use of antibiotics, the absence of colitis in rodent models in germ-free conditions, increased bacterial numbers in CD intestinal mucosa, and imbalances in microbial composition are examples (Table 5).

Twin-based studies have also contributed to demonstrating that intestinal bacteria play an important role in CD. Dicksved *et al.* [180] compared the intestinal microbial community of twins using terminal-restriction fragment length polymorphism (T-RFLP) and %G+C profiling. They observed that the similarity was higher mainly among non-IBD twins, but also among concordant IBD twins. The similarity did not hold for discordant IBD twins, suggesting that the microbiota of diseased individuals is different from their healthy twin, thus pointing to the importance of microbiota in CD.

Table 5. Evidence of bacterial implication in Crohn's disease

Year	Model	Finding	Ref.
1985	CD patients	Re-introduction of small-bowel effluent into the colon of patients with an ileostomy induces inflammation, whereas the re-introduction of a sterile ultra filtrated effluent does not.	[181]
1991	CD patients	Postsurgical exposure of terminal ileum to luminal contents results in increased inflammation, whereas diversion of fecal stream is associated with improvement.	[182]
1995	CD patients	Use of antibiotics delays recurrence after ileocecal resection.	[183]
1996	Rodent models	Bacteria are necessary for the development of chronic colitis.	[184]
1997	Humans	Intestinal bacterial replacement therapy, consisting of sterilisation and repopulation with normal intestinal microbiota, induces remission in refractory IBD patients.	[185]
1998	Human serum	Anti-*Saccharomyces cerevisiae* antibodies (ASCA) are strongly associated with IBDs.	[186]
1999	Rectal biopsies	Increased bacterial numbers within intestinal mucus of CD patients compared with controls. No correlation between bacterial concentration and disease severity.	[187]
2000	Feces	19.3% of active CD patients are positive for enterotoxigenic *Bacteroides fragilis* in feces.	[188]
2000	Biopsies	I2 sequence is associated with CD, especially Crohn's ileitis.	[189]
2002		I2 protein is present in *Pseudomonas fluorescens.*	[190]
2002		ASCA, anti-*E. coli* outer membrane protein C (ompC), and I2 are associated with CD.	[191]
2004		High levels of I2 and ompC are associated with different CD phenotypes - perforating and fibrostenosing disease respectively - and high risk of ileal surgery.	[192]
2002	Biopsies	IBD patients have higher concentrations of mucosa-associated bacteria than controls. In controls, intestinal mucosa is nearly sterile after removal of mucus.	[193] [194]
2002	Biopsies	Bacterial penetration is higher in CD patients than controls.	[194]
2002-to date		Alteration of microbial composition in IBD, resulting in reduction of diversity and dysbiosis*. Extended explanation in Section 4.2.1	[195, 196]
1978-to date		Detection of bacterial pathogens such as *Mycobacterium avium* sbsp. *paratuberculosis*, *Klebsiella pneumoniae*, *Listeria* spp., *E. coli*, and *Streptococcus* spp. in CD patients. Extended explanation in Section 4.3	[197-202]

4.2 Enteric microbiota in Crohn's disease

The number of studies analysing the composition of enteric microbiota associated with CD and other IBDs have significantly increased in the last few years. Due to the extreme complexity of the gut ecosystem and the anaerobic nature of its inhabitants, culture-based techniques have played a limited role in expanding knowledge in this field. In this section we summarise the results of those works that have principally used culture-independent techniques to describe the microbial community of CD patients.

Kleessen *et al.* [194] observed that bacteria were localised on the mucosal surface and in the submucosal tissue of IBD patients, but were absent in muscularis propria. Bacterial invasion was pronounced in areas with erosion and crypt abscesses*, and bacterial cell clusters or single cells could appear intracellularly in epithelial cells or phagocytes of lamina propria. As is shown in Table 5, IBD patients exhibit increased numbers of mucosa-associated bacteria compared with healthy subjects [187]. In particular, the highest counts were found in the ileum [187, 193, 194]. There is some controversy regarding inflamed and non-inflamed mucosa-associated microbiota. Swidsinski *et al.* [193] observed that in non-inflamed mucosa bacterial concentration is higher, whereas more recently Bibiloni *et al.* [203] and Vasquez *et al.* [204] revealed that the bacterial counts and composition of inflamed and non-inflamed mucosa of CD patients were similar.

Similar bacterial composition has also been reported for inflamed and non-inflamed mucosa [203, 205, 206]. In summary, the composition of the intestinal microbiota of CD patients has been described as: i) stable along the GI tract – at least from the ileum to the rectum, ii) host-specific, iii) differentiable from healthy subjects and patients suffering from UC, iv) less diverse than in controls, v) heterogeneous in comparison to healthy subjects, which have more bacterial groups in common, and vi) apparently unstable depending on the disease activity state [203, 205-210].

4.2.1 Dysbiosis as a feature of Crohn's disease

An unbalanced microbial composition, with a decrease in protective bacteria and an increase in harmful microorganisms, called dysbiosis, has been extensively reported in CD patients regardless of sample type and methodological approach. In Table 6, the taxons that have been most frequently found associated with CD patients and healthy subjects are listed. In summary, CD patients are

characterised by a decrease in *Firmicutes* [205, 207, 211-217], in particular *Clostridium* clusters XIVa and IV, and an increase in facultative anaerobic bacteria, mainly members of *Enterobacteriaceae* [204, 205, 207, 211-215, 217-219].

Controversial results have been obtained for the *Bacteroides-Prevotella* and *Bifidobacteria* groups. Several authors have associated *Bacteroides*, a common inhabitant of the intestinal mucosa, with CD [203, 204, 211, 212, 216] and others with healthy subjects [210, 213-215, 217, 218, 220].

A similar situation occurs with some members of *Clostridiales*. In general, this group is associated with healthy subjects; however, some authors have found certain species or groups linked to CD [204, 205, 214].

The case of *Bifidobacterium* is quite different. There is a tendency towards *Bifidobacterium* reduction in CD and other IBD patients, but it is not very noticeable. As a consequence, some authors find these differences significant [215, 221] whereas others do not [211, 217].

Kuehbacher *et al.* 2008 [222] have recently associated TM7 bacterial phylotypes with IBD. TM7 bacteria had already been linked to mucosal inflammation of the oral cavity. Kuehbacher *et al.* suggest that this group of bacteria could play a role in modulating the intestinal mucosa towards a "pro-inflammatory" microbiota.

Table 6. Imbalances detected in tissue-associated CD microbiota (** a selection of studies based on feces is also included). Microorganisms classified in different taxonomic levels have been associated with CD patients or healthy controls by their major prevalence or abundance in each group of patients as observed using different methodological approaches.

Reference	Method and sample size	Taxon associated with CD	Taxon associated with C
Prindiville *et al.* 2004 [214]	*Cloning* (739 clones) 41 Biopsies (31 from CD patients, 10 from C subjects)	*Enterococcus faecium* Facultative bacteria *Ruminococcus gnavus*	*Bacteroides fragilis*-like clone *Ruminococcus gnavus*-like clone *Prevotella nigrescens*-like clone *Clostridium leptum* subgroup
Swidsinski *et al.* 2005 [216]	*FISH* 20 CD, 20 UC 20 C, 20 IBS, 20 SLC	*Bacteroides*, especially *Bacteroides fragilis*	*Eubacterum rectale-Clostridium coccoides* *Faecalibacterium prausnitzii*
Manichanh *et al.* 2006 [213]	*Metagenomic analysis* (1190 clones, 125 OTUs) Healthy library from 6 C and I-CD patients' library from 6 CD.	"Uncultured *Porphyromonadaceae*" *Prevotella* subgroup	*Clostridium leptum* subgroup *Bacteroides fragilis*
Conte *et al.* 2006 [220]	*Culture, PCR, and RTi-PCR* Paediatric subjects: 12 CD, 7 UC, 6 IdC, 7 C	Aerobes and facultative anaerobes such as *E. coli*, *Klebsiella* spp, and *Proteus* spp.	*Bacteroides thetaiotaomicron* *Bacteroides vulgatus*

Table 6. Imbalances detected in tissue-associated CD microbiota (*Continued*)

Bibiloni *et al.* 2006 [203]	*DGGE, PCR, RTi-PCR and cloning* 15 UC, 12 C and 20 CD newly diagnosed patients	Unclassified *Bacteroidetes* Unclassified *Verrucomicrobia*	*Firmicutes* *Prevotella*
Gophna *et al.* 2006 [212]	*Cloning* (3305 clones, 150 OTUs)	*Proteobacteria*, especially *E. coli*, *Acinetobacter junii*, and *K. pneumoniae* *Fusobacteria* *Bacteroidetes*, principally *Bacteroides vulgatus* and *Bacterides fragilis*	*Firmicutes*, especially *Clostridia*
Scanlan *et al.* 2006 [210]	*DGGE* 16 CD (11 in remission), 6 C		*Clostridium leptum* subgroup *Bacteroides fragilis*
Martinez-Medina *et al.* 2006 [205]	*DGGE* 19 CD, 15 C, 2 UC, and 1 IC	*Enterobacteriaceae,* especially *Escherichia coli* *Clostridium coccoides* (subcluster XIVa) Uncultured *Ruminococcus torques*-like bacterium	*Firmcutes*, especially *Faecalibacterium prausnitzii*
Baumgart *et al.* 2007 [207]	*Cloning* (616 clones) 7 I-CD, 6 C-CD, 7 C	*Escherichia coli*	*Lachnospiraceae* *Clostridiales*, especially *Faecalibacteria* and *Subdoligranula*
Frank *et al.* 2007 [218]	*Cloning* (15172 clones) 192 subjects (CD/UC/C)	*Proteobacteria* (alpha, beta and gamma) *Actinobacteria*	*Bacteroidetes* *Lachnospiraceae* (*Clostridium* clusters XIVa and IV)
Dicksved *et al.* 2008 [180]	*T-RFLP, %G+C profiling and targeted-* Bacteroides *detection* 10 monozygotic twins (of which, 6 concordant for CD) and 8 healthy twins	Within *Bacteroides*: *Bacteroides ovatus* *Bacteroides vulgatus*	Within *Bacteroides*: *Bacteroides uniformis*
Andoh *et al.* 2008 [211]**	*T-RFLP* 34 CD (22 IC-CD, 8 C-CD, 4 I-CD), and 30 C	*Bacteroides* *Enterobacteriaceae*	*Clostridium* cluster IV, cluster XI, and subcluster XIVa
Kuehbacher *et al.* 2008 [222]	*Cloning TM7* 42 CD, 31 UC, 33 C	TM7 bacteria	
Swidsinski *et al.* 2008 [215]**	*FISH* 82 CD, 105 UC, 17 IdC, 268 non-IBD	*Enterobacteriaceae*	*Eubacterium rectale / C. coccoides* group *Bacteroidaceae* *F. prausnitzii*
Takaishi *et al.* 2008 [217]**	*FISH, RTi-PCR, culture* 23 CD (15 IC-CD, 8 C-CD), 73 UC, 65 C	Facultative anaerobes such as *Lactobacillus* and *Enterococcus*	*Bacteroides fragilis* *C. coccoides* group *C. leptum* subgroup *Atopobium*
Willing *et al.* 2009 [219]	*T-RFLP and RTi-PCR* 10 monozygotic twins (of which, 4 concordant for CD) and 6 healthy twins	*Escherichia coli* (specifically I-CD)	*Faecalibacterium prausnitzii* (in C and C-CD)

Abbreviations: C, control subjects; UC, Ulcerative colitis; CD, Crohn's disease; I-CD, Crohn's ileitis; IC-CD, Ileocolonic disease; C-CD, Crohn's colitis; IC, Ischemic colitis; IBS, Irritable Bowel Syndrome; SLC, Self-limiting colitis; I, inflamed; NI, non-inflamed; OTU, operational taxonomic unit.

Very recently, Ott *et al.* [223] have published a study describing the fungal population inhabiting the intestinal mucosa. CD patients showed an increased richness and diversity of fungi, with *Rhodutorula mucilaginosa*, *Exidiopsis calcea*, *Cyrtandra ootensis*, *C. parapsilosis*, *C. dubliensis*, *Trametes vesicolor*, *Filobasidium globisporum*, and *Saccharomyces cerevisiae* being frequent in CD and representing a shift in composition. However, fungi make up only 0.02% of the mucosa-associated flora.

4.3 Pathogens proposed as possible causative agents

Several enteric infections can mimic clinicopathological features of CD. *Campylobacter*, *Salmonella*, *Shigella*, *Yersinia*, *Escherichia coli* O157:H7, *Chlamydia trachomatis*, *Mycobacterium tuberculosis*, *Actinomyces* spp., *Entamoeba histolytica*, *Giardia lamblia*, and *Clostridium difficile* are pathogens that could drive infections causing confusion ([224] and references therein). Nevertheless, only two main microorganisms, *Mycobacterium avium* subspecies *paratuberculosis* and *Escherichia coli*, are being actively investigated at present.

4.3.1 *Listeria* spp. and *Yersinia* spp.: "the cold chain hypothesis"

The cold chain hypothesis is rooted in epidemiological data on the assessment of aspects of western life such as diet, hygiene, and refrigeration, as environmental risk factors for CD [34]. Based on the fact that cold-chain development paralleled the growth of CD during the 20th century, this hypothesis suggests that pathogenic psychotropic bacteria such as *Listeria* spp. and *Yersinia* spp. could contribute to the disease.

Yersinia is an enteric pathogen that causes clinical symptoms similar to CD, including the presence of granulomas*, microabscesses, ulceration and the activation of the NF-κB pathway with proinflammatory cytokine release. Pathogenic *Yersinia* DNA was detected in bowel and mesenteric lymph nodes from patients with CD by Lamps *et al.* [199], but in many other studies this bacterium was not found.

The gram-positive bacterium *Listeria monocytogenes* can also cause clinical symptoms characteristic of CD, including granulomas*. Although it was detected in 75% of CD patients by immunocytochemistry in a work that included 16 patients from the USA and France [200], many subsequent studies did not corroborate these results [197, 216, 225, 226]. In fact, several reports have implicated *Listeria* spp. in opportunistic infections after administration of

infliximab, thus indicating that this bacterium could be a consequence of the disease rather than the cause [227].

4.3.2 *Mycobacterium avium* subspecies *paratuberculosis* (MAP): similarities with Johne's disease

Mycobacteria have for a long time been suspected of being implicated in the pathogenesis of CD. The similarities in clinical manifestations of CD with Johne's disease in ruminants, which is caused by *Mycobacterium avium* subspecies *paratuberculosis* (MAP), lies behind this hypothesis. MAP has the ability to survive and replicate within macrophages. This resistance to phagocytosis by macrophages could explain the formation of granulomas which are present in certain CD phenotypes. MAP was firstly isolated from CD tissues in 1984 [228] and has been investigated in numerous studies by means of culturing, polymerase chain reaction (PCR), FISH and serology, but the results obtained have been contradictory. Although the majority of studies report a higher prevalence of MAP in CD patients, its occurrence can vary from 0 to 100% of patients sampled, and it can occasionally be detected in healthy subjects as well [193, 229-235]. Furthermore, treatment with triple antimycobacterial antibiotics failed to show a sustained response in a two-year prospective study [236]. Feller *et al.* [237] have recently published a systematic review and meta-analysis based on 28 case-control studies. They conclude that a causal role for MAP in the etiology of CD can be neither confirmed nor excluded with confidence, and that additional studies and systematic reviews are therefore needed.

4.3.3 *Helicobacter* spp.

Helicobacter spp. have been found more frequently in CD patients than in controls, especially in those patients with a UC-like phenotype [238, 239]. This observation, in conjunction with recent research based on animal models [240], points to this bacterial genus having a putative role in CD, although further studies are needed.

4.3.4 Yeasts

The presence of anti-*Saccharomyces cerevisiae* mannan antibodies (ASCAs) is closely associated with CD. This serological marker is very powerful because it is found in 50-60% of CD patients and 20-30% of their healthy relatives, whereas only 3-7% of the healthy control population gives positive results [186].

Although several microorganisms that are present in the gut can stimulate the expression of ASCAs, clinical and experimental evidence indicates *Candida albicans* is the major one implicated [241, 242]. *C. albicans* is a frequent intestinal commensal that can occasionally act as an opportunistic pathogen. Increased numbers of *C. albicans* have been found in CD-healthy relatives, correlating with an increased prevalence of ASCAs. In CD patients, the concentration of ASCAs is stable over time and is associated with severe phenotypes and ileal location. The hypothesis concerns a genetic defect in CD patients which confers susceptibility to ASCA-inducing microorganisms. Some studies have correlated *CARD15* mutations [243-246] and, more recently, mutations in the mannan-binding lectin gene with ASCA-positivity [247]. However, other studies have reported contrasting results [248-250].

5 *Escherichia coli* and Crohn's disease

5.1 Evidence of *E. coli* contribution to Crohn's disease pathogenesis

Increased levels of *Enterobacteriaceae* have been reported for Crohn's disease for four decades [251, 252]. Studies based on a variety of methods, such as culture [253], dot-blot hybridisation [100] and fluorescence *in situ* hybridisation [194], agree that *Enterobacteriaceae* are more abundant in CD patients than in controls. Further studies focusing on *Escherichia coli* indicate that this bacterium is the principal enterobacterial species that increases in the CD population [207, 212, 216, 254, 255]. In addition, high levels of antibodies against *E. coli* outer membrane protein C (OmpC) have also been observed in CD patients, correlating with disease progression, longer duration, and greater need for surgery [191, 192, 256].

The direct implication of *E. coli* in Crohn's disease pathogenesis is still not clear. However, several observations indicate that this bacterium can play an important role. For example, after an ileocolonic resection, high counts of *E. coli* (and also *Bacteroides* spp.) predispose to an early recurrence [257]. Another indicator is that *Enterobacteriaceae* account for 24% of penetrated bacteria in CD patients [194]. Moreover, *E. coli* has been detected more frequently in the intestinal serosa, mesenteric lymph nodes, and granulomas* of CD patients (27%, 33% and 80% respectively) than of controls [258-260].

Very recently, increasing attention has been paid to investigating the pathogenic properties of *E. coli* strains inhabiting the gut of CD patients and healthy controls. However, further studies are still necessary to achieve a complete characterisation. In a study in 1997, Schultz *et al.* [261] were already aiming to describe the presence of virulence factors, but they were not able to find differences between *E. coli* strains isolated from CD patients and controls. In contrast, more recent works indicate that there are differences of phylogenetic origin – with B2+D phylogroups being more frequently found in CD patients [262] – and phenotypic virulence properties – with *E. coli* having adhesion and invasion properties being more prevalent in CD patients [263].

5.2 Pathogenic *E. coli* and the recently described adherent-invasive *Escherichia coli* pathovar

Non-pathogenic *E. coli* strains commonly colonise the human intestinal tract and interact with the host in a mutualistic relationship. However, in immunosuppressed hosts or when the gastrointestinal barrier is broken, even non-pathogenic *E. coli* strains can cause infection. Moreover, a number of strains have evolved the ability to cause a broad range of human diseases. Pathogenic *E. coli* are classified in two main groups depending on the type of infection that they cause. One group, named extraintestinal pathogenic *E. coli* (ExPEC), enclose those strains causing infections such as urinary tract infections (UTI), sepsis, and meningitis. The other group includes those strains that cause intestinal infections and are called diarrhoeagenic *E. coli* (DEC). Among the intestinal pathogens, there are six well-described pathotypes: enteropathogenic *E. coli* (EPEC), enterohemorrhagic *E. coli* (EHEC), enterotoxigenic *E. coli* (ETEC), enteroaggregative *E. coli* (EAEC), enteroinvasive *E. coli* (EIEC), and diffusely adherent *E. coli* (DAEC) (Figure 10). Uropathogenic *E. coli* (UPEC), which is the main cause of UTIs, and meningitis-associated *E. coli* (MNEC), which is responsible for meningitis and sepsis, are the two pathotypes described to date from the ExPEC group. An additional animal pathotype, known as avian pathogenic *E. coli* (APEC), has been described for poultry and causes extraintestinal infections. New emerging pathotypes are currently being proposed. Among them, adherent-invasive *E. coli* (AIEC) has been associated with Crohn's disease (further information in Section 5.2.9).

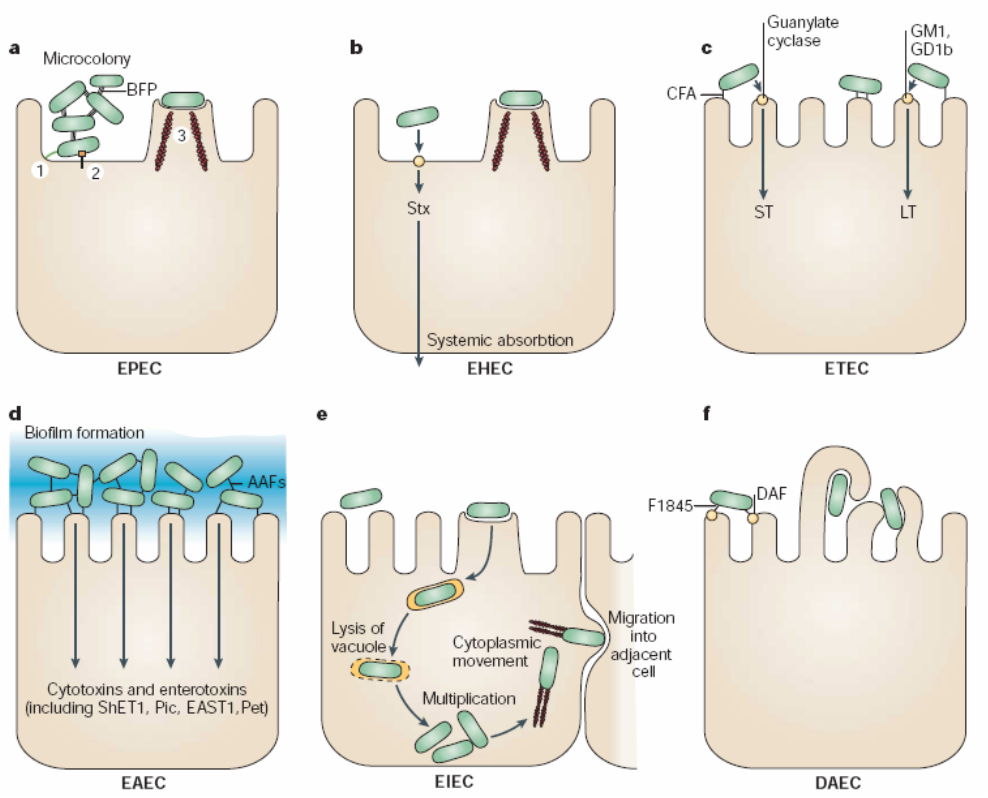


Figure 10. Pathogenic mechanisms of diarrhoeagenic *E. coli.* Source: Nataro and Kaper, 2004 [264]. Pathogenicity models are based on *in vitro* studies.

a 1) Formation of microcolonies that adhere to small bowel enterocytes, 2) Protein translocation by type III secretion, and 3) pedestal formation. Downstream effects are: destruction of microvillar architecture and loss of absorptive surface area, increased permeability, active ion secretion, and intestinal inflammation.

b EHEC adhere specifically to enterocytes of the colon and promote changes in their cytoskeleton, as with EPEC, and finally forming a pedestal. In addition, EHEC synthesise the Stx toxin, absorption of which leads generally to diarrhoea or other complications.

c ETEC adhere to small bowel enterocytes by means of CFA fimbrial/fribrillar proteins. ETEC then synthesise LT and/or ST enterotoxins which give rise to a variety of mechanisms that lead to intestinal secretion and decreased intestinal absorption, thus provoking watery diarrhoea.

d EAEC adhere to enterocytes and form a thick biofilm, give rise to small cytoskeletal changes, and synthesise enterotoxins and cytotoxins that cause intestinal secretion.

e EIEC penetrate intestinal cells, lyse endocytic vacuoles, replicate intracellularly, move through cellular cytoplasm by nucleating actin filaments and finally migrate to adjacent cells. In addition, they trigger macrophage apoptosis and release of cytokines.

f DAEC adhere to small bowel enterocytes, giving rise to cellular changes. Long finger-like cellular projections wrap around the adherent bacteria. Seventy-five percent of DAEC strains produce the F1845 fimbrial adhesin or other Dr adhesin, which bind with DAF, a cell surface glycoprotein.

Abbreviations: AAF, aggregative adherence fimbriae; BFP, bundle-forming pilus; CFA, colonisation factor antigen; DAF, decay-accelerating factor; EAST1, enteroaggregative *E. coli* ST1; LT, heat-labile enterotoxin; ShET1, *Shigella* enterotoxin; ST, heat-stable enterotoxin; Stx, Shiga toxin.

Pathogenic *E. coli*, like many other mucosal pathogens, share a common progression of pathogenesis which consists of i) colonisation of a mucosal site, ii) evasion of host defences, iii) multiplication, and iv) host damage [265]. In addition, EIEC has the ability to invade and replicate within epithelial cells and macrophages. Each pathotype possesses specific virulence factors and even among the strains of a certain pathotype there are different virulence gene sets. Among these virulence factors, molecules or structures responsible for adhesion/colonisation – for instance, fimbriae (also called pili), fibrillae, and other afimbrial adhesins such as outer membrane proteins – factors that suit colonisation – such as siderophores* and other molecules implicated in iron uptake, and capsules – can be found in pathogenic *E. coli*. A wide range of toxins and other effector proteins that have an influence on a wide variety of eukaryotic processes, including cell lysis, cytoskeleton destruction or alteration, and apoptosis, can also be found. Additional information concerning virulence genes known for *E. coli* can be found in the Virulence Factor Database (VFDB) created by the State Key Laboratory for Molecular Virology and Genetic Engineering, Beijing, China [266, 267]. The VFDB has recently been upgraded (2008) to make it a platform for further study of comparative pathogenomics.

5.2.1 Enteropathogenic *E. coli* (EPEC)

EPEC was the first pathotype to be described and, in developing countries, is still an important pathogen that can cause fatal infant diarrhoea [264].

The main characteristics of EPEC are:

- Intestinal histopathology known as the ‘attaching and effacing (A/E) lesion’. The bacteria attach tightly to the epithelium and promote changes in the cytoskeleton of eukaryotic cells, which results in the effacement of microvilli and the formation of pedestal-like structures on which the bacteria perch (Figure 10a). This characteristic is encoded by different genes located in the ‘locus of enterocyte effacement’ (LEE) of a pathogenicity island (PAI) which contains information on other virulence factors. EHEC strains also exhibit A/E histopathology.
- The presence of the EAF plasmid (EPEC adherence factor), which encodes for type IV pilus (also called bundle-forming pilus (BFP)), and the *per* locus (plasmid encoded regulator). BFP mediates bacteria-bacteria and bacteria-enterocyte adhesion in a characteristic pattern termed ‘localised adherence’ (LA), forming clusters or microcolonies on the surface of host cells. An LA pattern is a feature of EPEC strains and, therefore, has been used widely as a

diagnostic tool. In turn, the *per* locus regulates the *bfp* operon and most genes in the LEE.

- Only some EPEC produce enterotoxins.
- Major virulence factors in EPEC:

▶ Adherence	● *bfp* (Bundle-forming pilus) ● *eae* (Intimin) ● LifA (Lymphostatin) ● Paa (Porcine attaching-effacing associated protein)
▶ Protease	● EspC
▶ Regulation	● *Ler* ● *Per*
▶ Secretion system	● TTSS (Type III secretion system)
▶ Toxin	● CDT (Cytolethal distending toxin) ● EAST1
▶ Type III translocated protein	● Cif (Cycle inhibiting factor) ● EspA, B, D, F, G, H (Effector secretion proteins) ● Map (Mitochondria-associated protein) ● NleA/EspI (Non-locus-of-enterocyte-effacement-encoded effector A) ● NleC, D (Non-locus-of-enterocyte-effacement-encoded effector C, D) ● Tir (Translocated intimin receptor)
▶ Pathogenicity islands	● EspC island ● LEE (EPEC)

5.2.2 Enterohemorrhagic *E. coli* (EHEC)

EHEC cause bloody diarrhoea, non-bloody diarrhoea and the haemolytic uremic syndrome. Strains of the O157:H7 serotype are the most EHEC pathogens in the USA, the United Kingdom, and Japan; however other specific serotypes, such as the O26 and O111 serogroups, are known to cause disease [265] in many other countries.

The main characteristics of EHEC are:

- Production of Shiga toxin (Stx), also known as verocytotoxin (VT) (Figure 10b). There are two subgroups, Stx1 and Stx2, which are 55% homologous. Stx induces apoptosis in intestinal epithelial cells. When these toxins travel via the bloodstream to the kidney, the damage to renal endothelial cells leads to inflammation and probably further haemolytic uremic syndrome.
- Most EHEC also contain a LEE pathogenicity island that encodes type III secretion systems and effector proteins homologous to those that produce EPEC strains. Those strains that produce Stx but lack a LEE pathogenicity island are called Shiga toxin-producing *E. coli* (STEC). The majority of STEC do not cause disease, indicating the importance of LEE. In fact, it is thought

that the EHEC pathovar evolved from an LEE-containing EPEC which acquired Stx-encoding bacteriophage.

- Major virulence factors in EHEC:

▶ Adherence	● ECP (*E. coli* common pilus) ● Efa-1/LifA (Lymphostatin) ● *eae* (Intimin) ● Paa (Porcine attaching-effacing associated protein) ● ToxB
▶ Iron uptake	● Chu (*E. coli* hemin uptake)
▶ Protease	● EspP ● StcE (Secreted protease of C1 esterase inhibitor from EHEC)
▶ Regulation	● *Ler*
▶ Secretion system	● TTSS (Type III secretion system)
▶ Toxin	● Hemolysin ● Stx (Shiga toxin)
▶ Type III translocated protein	● Cif (Cycle inhibiting factor) ● EspA, B, D, F, G, H (Effector secretion proteins) ● Map (Mitochondria-associated protein) ● NleA, C, D /EspI (Non-locus-of-enterocyte-effacement-encoded effector A, C, D) ● Tir (Translocated intimin receptor)
▶ Pathogenicity islands	● LEE (EHEC)

5.2.3 Enterotoxigenic *E. coli* (ETEC)

ETEC cause mild to severe watery diarrhoea. In developing countries, EPEC are important infectious agents for infant diarrhoea, whereas in developed countries they are the main cause of travellers' diarrhoea. EPEC colonise the ileal mucosa and secrete enterotoxins that promote intestinal secretion, thus leading to diarrhoea ([265] and references therein).

The main characteristics of ETEC are:

- Production of heat-labile enterotoxins (LTs) and/or heat-stable enterotoxins (STs). LTs are similar in function and structure to the cholera enterotoxin. LT-I is associated with human disease and LT-II with animal disease. LTs consist of a single *A* subunit responsible for enzyme activity and five *B* subunits which mediate the adhesion of the toxin to the enterocytes via GM1 and GD1b cell surface gangliosides (Figure 10c). LTs promote the increased secretion of Cl^- by secretory crypt cells through an adenylate cyclase-mediated signalling cascade. LTs can also stimulate prostaglandin synthesis and the enteric nervous system; all these effects give rise to increased secretion and decreased absorption. STs are single peptide proteins that comprise STa (only found in humans) and STb (associated with animals) disease types. The final result of

both toxins is increased intestinal secretion, but each ST type has its own mechanism.

- Colonisation is mediated by one or several colonisation factors (CFs) that are mainly fimbrial or fibrillar proteins. These factors are commonly designated CFA (colonisation factor antigen).
- ETEC strains typically possess multiple plasmids. The genes encoding CFs are generally found on a plasmid that also encodes ST and/or LT.
- Major virulence factors in ETEC:

► Adherence	● Adhesive fimbriae (21 different CFs, but around 75% of human ETEC express either CFA/I, CFA/II or CFA/IV).
► Toxin	● Heat-labile toxin (LT) ● Heat-stable toxin (ST)

5.2.4 Enteroaggregative *E. coli* (EAEC)

EAEC are the cause of persistent diarrhoea in infants and adults in developing and industrialised countries. No specific virulence factor has been described for all EAEC strains to date. In fact, this pathovar seems to contain a package of plasmid-borne and chromosomal virulence factors similar to those of other enteric pathogens. Furthermore, both pathogenic and non-pathogenic strains are included within this pathotype.

The main characteristics of EAEC are:

- Absence of LT and ST enterotoxin secretion.
- Autoaggregative (AA) adhesion to intestinal cells, in which bacteria adhere to each other in a 'stacked-brick' configuration (Figure 10d). Aggregative adherence fimbriae (AAFs) permit interbacterial adherence, which gives rise to biofilm formation and adherence to intestinal cells. However, not all the EAEC strains adhere by means of AAFs and other structures that may be involved in their adhesion have been described (e.g. dispersin, and EAEC flagellins).
- 'Typical EAEC' are considered to be those carrying the transcriptional activator AggR, which regulates the transcription of several EAEC virulence factors, whereas 'atypical EAEC' are those lacking AggR as proposed by Kaper and Nataro [265].
- Synthesis of enterotoxins and/or cytotoxins.

- Major virulence factors in EAEC:

▶ Adherence	● AAFs (Aggregative adherence fimbriae) ● Dispersin (anti-aggregation protein)
▶ Toxin	● EAST1 (EAEC heat-stable enterotoxin 1) ● Pet (plasmid-encoded enterotoxin) ● Pic (protein involved in intestinal colonisation) ● ShET1 (*Shigella* enterotoxin 1)

5.2.5 Enteroinvasive *E. coli* (EIEC)

EIEC are very close to *Shigella* spp., both genetically and pathogenically. They can be distinguished by a few minor biochemical features. Normally, EIEC cause watery diarrhoea that is indistinguishable from that due to infection by other DEC. However, EIEC can cause invasive inflammatory colitis and, occasionally, dysentery clinically indistinguishable from that caused by *Shigella*, spp ([265] and references therein).

The main characteristics of EIEC are:

- Intracellular pathogen. EIEC penetrate epithelial cells followed by the lysis of endocytic vacuole and further intracellular multiplication. They then move through the cytoplasm by nucleation of cellular actin and extend to adjacent cells (Figure 10e).
- Virulence genes are coded in a large plasmid also present in all *Shigella* species. One third of the plasmid contains insertion sequence elements which contribute to its evolution. The rest encode for a type III secretion system and an outer membrane protein involved in the nucleation of the cellular actin. Proteins secreted by the type III secretion system mediate epithelial signalling events, cytoskeletal rearrangements, cellular uptake, lysis of endocytic vacuole, and cellular pore formations. Additional virulence genes of EIEC and *Shigella* spp. that are plasmid- encoded, chromosomally-encoded and PAIs-encoded have been described.
- Major virulence factors in EIEC and *Shigella* spp.:

▶ Actin-based motility	● IcsA (VirG) (Intercellular spread)
▶ Endotoxin	● LPS (Serotype converting genes)
▶ Iron uptake	● Aerobactin
▶ Protease	● IcsP (SopA) ● Pic (Protein involved in intestinal colonisation) ● SigA (*Shigella* IgA-like protease homology)
▶ Secretion system	● TTSS (Type III secretion system)
▶ Toxin	● ShET1 and ShET2 (*Shigella* enterotoxins 1 and 2)
▶ Pathogenicity islands	● SHI-1, SHI-2, and SHI-3 ● SRL (*Shigella* resistance locus)

5.2.6 Diffusely adherent *E. coli* (DAEC)

DAEC have been implicated in infant diarrhoea, especially in children more than one year old [264, 268], as well as in many cases of UTIs ([269] and references therein).

The main characteristics of DAEC are:

- Distinctive diffuse pattern of adherence (DA) to HEp-2 cell monolayers (Figure 10f).
- Two subclasses of DAEC strains: diffusely adhering enteropathogenic *E. coli* (DA-EPEC) harbouring a LEE island, and DAECs expressing adhesins of the Afa/Dr family, which are responsible for UTIs.
- Major virulence factors in DAEC:

▶ Adherence	● Afa/Dr family including afimbrial adhesins AfaE-I, AfaE-III and Dr-II as well as the fimbrial Dr and F1845 adhesins
▶ Pathogenicity islands	● LEE

5.2.7 Uropathogenic *E. coli* (UPEC)

UPEC are the most common cause of community-acquired (70-95%) and nosocomial (50%) urinary tract infections [270]. UPEC generally colonise the bladder, causing cystitis, but they can also ascend into the kidneys and cause pyelonephritis. UPEC bind host tissue and invade urothelial cells, where they are trafficked into membrane-bound acidic compartments similar to lysosomes [271]. UPEC can replicate intracellularly, forming biofilm-like communities. Infected bladder cells are primed to exfoliate, leaving immature bladder cells exposed and more susceptible to infection. Late endosome-like compartments of immature bladder cells contain UPEC in a quiescent state that may serve as a reservoir for recurrent UTIs [272].

UPEC harbour more virulence associated genes, usually encoded in PAIs, than commensal *E. coli* isolates. The principal virulence factors of UPEC encode for capsule antigens, iron acquisition systems, adhesins, and secreted toxins. Despite the fact that DEC pathotypes also harbour numerous virulence genes in PAIs, their nature is different from those of UPEC.

Major virulence factors in UPEC are:

▶ Adherence	● Dr adhesins ● F1C fimbriae (*foc*) ● P fimbriae (*pap*) ● S fimbriae (*sfa*) ● Type I fimbriae (*fim*)
▶ Iron uptake	● Enterobactin siderophores ● Aerobactin siderophores ● IroN (Salmochelin siderophore) ● Chu (*E. coli* hemin uptake) ● Sit (Iron/manganese transport) ● fyuA (Yersiniabactin siderophore)
▶ Protease	● Pic (Protein involved in intestinal colonisation) ● Sat (Secreted autotransporter toxin) ● Tsh (Temperature sensitive hemagglutinin)
▶ Toxin	● CNF-1 (Cytotoxic necrotising factor 1) ● Hemolysin
▶ Pathogenicity islands	● PAI I_{536}, PAI I_{CFT073}, PAI II_{536}, PAI III_{536}

The above-mentioned virulence factors can be present at different frequencies among various subgroups of UPEC. In addition, a single UPEC strain can accumulate several virulence genes associated with adherence or iron uptake. For example, UPEC strains can carry more than 10 fimbrial gene clusters, and the environmental conditions determine the expression of one pilus or another. This process is known as 'phase variation'. Among adhesive organelles, type 1 pilli are highly conserved and common among all *E. coli*, both the pathogenic and commensal isolates. FimH adhesin is known to play an important role in bacterial adherence to and invasion of host cells, and to contribute to intracellular biofilm formation in UPEC [273, 274]. With regard to the secreted toxins, instead of the type III secretion system, UPEC often use the type I, for which α-haemolysin is the prototypical secreted toxin, and type V secretion systems, which includes toxins known as autotransporters. This arsenal of toxins may promote cell lysis, apoptosis, inflammatory responses, and cytopathic effects such as vacuolation and swelling.

5.2.8 Meningitis/sepsis-associated *E. coli* (MNEC)

MNEC is the main cause of gram-negative neonatal meningitis, with a high mortality rate (15-40%) and severe neurological defects in most survivors. The transmission of NMEC to the central nervous system occurs haematogenously. Increased levels of bacteraemia confer a risk of meningitis ([265] and references therein).

A restricted number of serotypes represent the MNEC pathovar, with the K1 capsule type being present in 80% of the strains, and O18:K1:H7 the prototypic MNEC. The K1 capsule confers invasiveness, modulates trafficking of *E. coli*-

containing vacuoles and enhances intracellular bacterial survival in brain microvascular endothelial cells.

Major virulence factors in MNEC are:

▶ Adherence	● S fimbriae (*sfa*)
▶ Invasion	● AslA ● Ibe A, B and C (Invasion of brain endothelial cells) ● K1 capsule ● OmpA (Outer membrane protein A) ● TraJ
▶ Toxin	● CNF-1 (Cytotoxic necrotising factor 1)

5.2.9 Adherent-invasive *E. coli* (AIEC), an emergent *E. coli* pathotype associated with Crohn's disease

Although considerable effort has been expended in studying the *E. coli* populations of CD patients, none of the above intestinal pathogenic *E. coli* have been found to be associated with the disease. Nevertheless, in 1998 Darfeuille-Michaud *et al.* [263] observed that adherent *E. coli* strains colonised the ileal mucosa of CD patients more frequently than in control groups. Further studies indicated that these adherent *E. coli* strains had other pathogenic traits which differentiated them from the rest of the *E. coli* pathovars, and they therefore designated a potentially new *E. coli* pathovar named adherent-invasive *E. coli* (AIEC) [275].

The main characteristics of AIEC are:

- Ability to adhere to and to invade intestinal epithelial cells. The invasion occurs through a macropinocytosis-like process which is dependent on actin microfilaments and microtubules recruitment [276]. In addition, the AIEC type 1 pilus induces the elongation of the epithelial cell membrane surrounding the bacteria at the point of contact. The interaction between AIEC and cultured intestinal epithelial cells leads to an upregulation of IL-8 and CCL20 expression, giving rise to the transmigration of polymorphonuclear leukocytes and dendritic cells [277] and provoking some changes in the epithelial cell structure, leading to reduced barrier resistance [278].
- Ability to survive and to replicate extensively within macrophages without triggering host cell death and inducing the release of TNF-α [279]. Moreover, one study reports the formation of granuloma*-like cell aggregates after the infection of an *in vitro* model of human granuloma by AIEC [280].
- Absence of known invasive determinants.

No specific virulence factors characteristic of the AIEC pathovar are known to date. However, some genes and regulatory processes have been implicated in the pathogenesis of the prototypic AIEC strain LF82. **Type 1 pili** play an important role in AIEC colonisation, being responsible for adhesion to the carcinoembryonic antigen-related cell adhesion molecule 6 (CEACAM6) expressed on the apical side of enterocytes. Interestingly, whereas a type 1 pilus is necessary for AIEC to invade host cells, the expression of type 1 pili in non-pathogenic *E. coli* K12 is not sufficient to confer invasiveness [276]. Moreover, *fim* gene sequencing revealed that LF82 produce a variant of type 1 pili in comparison to those of *E. coli* K12 [276]. **Flagella** also play a direct role in the adhesion-invasion processes of LF82 via motility and an indirect role via regulating type 1 pili and other unknown factors [281]. Lipoprotein **NlpI** is thought to be involved in the regulatory pathway of flagella, type 1 pili and other unknown virulence genes [282]. In turn, lipoprotein **YfgL** was reported to be required for the invasiveness of LF82, independently of the type 1 pilus and flagellum but in association with the release of outer membrane vesicles [283]. Recently, Rohlion *et al.* [284] proposed a model in which the outer membrane protein C (**OmpC**), a porin regulated by EnvZ/OmpR, is involved in the adherence-invasiveness of AIEC involving type 1 pili, flagella and other unknown virulence factors. Interestingly, type 1 pili and flagella encoding gene regulation were opposite in AIEC LF82 and non-pathogenic *E. coli* K12. Finally, the oxidoreductase **DsbA** [285] and **HtrA** [286] stress protein have been implicated in the ability of LF82 to survive and replicate within macrophages. Both genes are present in *E. coli* K12 as well, but a differential regulation dependent on LF82 genetic background marks the difference.

A model explaining the putative mechanism of pathogenesis of AIEC is shown in Figure 11. AIEC adhere via type 1 pili to the CEACAM6 receptor, which is abnormally expressed in epithelial cells from CD patients [287]. Once adhered, flagellin recognition by TLR5 induces IL-8 cytokine secretion, leading to the transmigration of polymorphonuclear leukocytes and dendritic cells. AIEC then cross the mucosal barrier thanks to their ability to invade epithelial cells. Macrophages engulf translocated bacteria, which resist macrophage phagocytosis and replicate within large vacuoles, inducing TNF-α secretion. The presence of intracellular AIEC plus the secretion of TNF-α and interferon (IFN-γ) promote the overexpression of the CEACAM6 receptor, thus enhancing AIEC colonisation and intestinal inflammation.

The AIEC pathovar has achieved greater significance as a potential etiological agent in CD since a number of independent studies from several countries –

France [288], the United Kingdom [289], and the USA [207, 278] – reported a higher prevalence of intramucosal or mucosa-associated *E. coli* with invasive properties in CD patients than in control subjects.

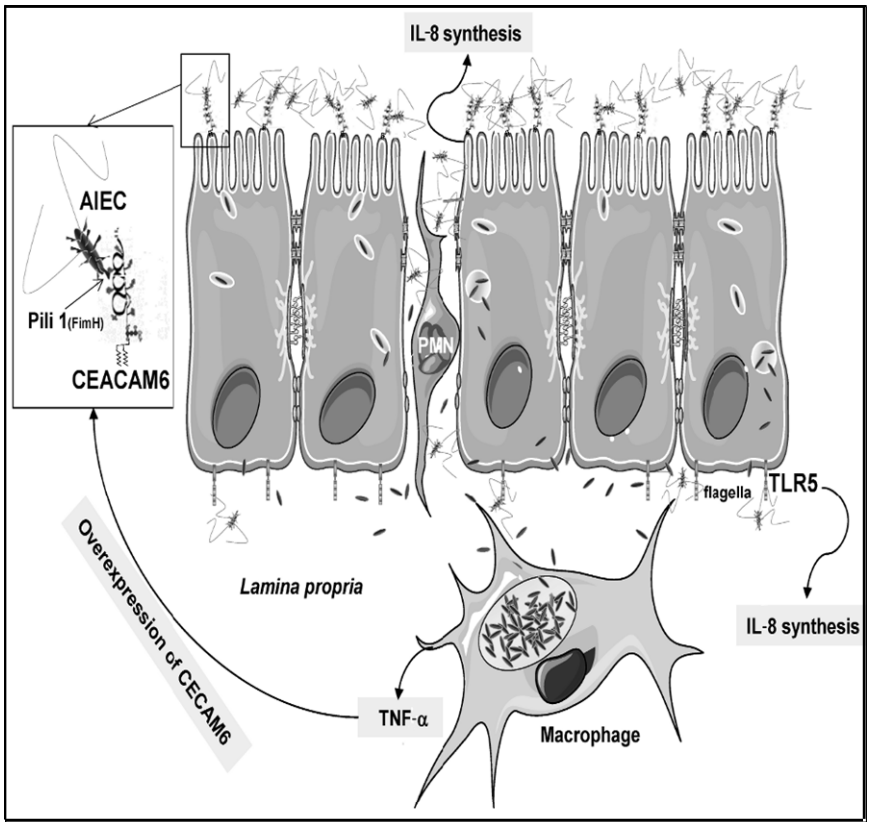


Figure 11. Putative pathogenic mechanism of adherent-invasive *E. coli* (AIEC). Source: Glasser and Darfeuille-Michaud, 2008 [290].

Objectives

Chapter 1: Description of the bacterial community in Crohn's disease patients and the search for compositional differences with respect to control subjects

By the time this work began, evidence already existed that intestinal microbiota was in some way implicated in Crohn's disease (CD). However, whether bacteria played a direct or indirect role in the disease was still controversial because no pathogenic agent had been demonstrated as being the cause. Some researchers suggested that the mainly culture-based methodology used up until then and that fecal samples were being used rather than biopsies, were limiting factors that might have been misleading researchers looking for a putative etiologic agent. For that reason, the **main goal** of the first part of our study was to use a molecular-based approach to analyse the enteric microbial community in CD patients, focusing on those bacteria close to the intestinal mucosa and making comparisons with control subjects and patients suffering from other inflammatory bowel diseases (IBD). The specific objectives were:

- To find bacterial species that characterise Crohn's disease and that could play an etiological role in the disease due to their proximity to host mucosal cells.
- To search for possible variations in the mucosal bacterial community along the gastrointestinal tract, in inflamed/non-inflamed areas and in ulcerated/non-ulcerated mucosa.
- To determine whether or not the localisation and severity of the disease, type of medication, resection and other clinical data of CD patients had an effect on resident intestinal microbiota.

Chapter 2: Characterisation of *Escherichia coli* populations associated to the intestinal mucosa of Crohn's disease patients and control subjects

Several bacterial species and phylogenetic groups were found to be associated with the intestinal mucosa of CD patients during the first part of the study. Phylotypes belonging to γ-*Proteobacteria*, especially to the family *Enterobacteriaceae*, were found to be more prevalent in CD patients than in control groups. Among them, *Escherichia coli* was the species most frequently detected. At the same time, there was a growing body of evidence from other researchers indicating that this bacterium could be playing an important role in CD, in particular a very recently described pathovar: adherent-invasive *E. coli* (AIEC). For that reason, we focused on the *E. coli* population and designed a second study with the following aims:

- To describe and compare the *E. coli* populations associated with the intestinal mucosa of CD patients and control subjects in search of CD-specific clones.
- To compare the richness, abundance and diversity of *E. coli* subtypes and to look for ecologic imbalances at a subspecies level.
- To compare the pathogenic features of *E. coli* subtypes isolated from CD patients and non-IBD controls, by characterising the strains phenotypically and genotypically.
- To determine the prevalence of AIEC in our set of patients and compare this with studies from other countries in order to test the hypothesis that AIEC is implicated in CD.

Chapter 3: Characterisation of adherent-invasive *E. coli* strains isolated from the intestinal mucosa of Crohn's disease patients and control subjects

During the second part of the study, we observed that the *E. coli* populations from CD patients and control subjects differed only in quantity. No specific clones were detected in CD patients; there were no differences in terms of the richness and diversity of *E. coli* subtypes; no virulence genes were found to be more frequent in the *E. coli* from CD patients and neither was any difference observed in the phylogenetic origin of the strains. In contrast, when we identified AIEC over the entire *E. coli* collection gathered in this work, we found a correlation between CD patients and a higher prevalence, abundance and richness of AIEC strains. Finally, AIEC strains have been reported to be similar to other pathogenic *E. coli* responsible for extraintestinal infections. Therefore, for the third part of our study, we focused on a more detailed characterisation of the AIEC and non-AIEC strains obtained in the second part. The main aims of this chapter were:

- To characterise the AIEC collection by identifying additional virulence traits, including virulence-associated genes and biofilm formation capacity.
- To determine whether a collection of the main urinary pathogenic *E. coli* and sepsis-meningitis pathogenic *E. coli* strains share the phenotypical traits that characterise the AIEC pathovar.
- To search for a common phylogenetic origin between those intestinal and extraintestinal AIEC strains.

Results

Chapter 1

Description of the bacterial community in Crohn's disease patients and the search for compositional changes with respect to control subjects

Introduction

Crohn's disease (CD) is a chronic disorder with chronic morbidity characterized by patchy inflammation of the gastrointestinal tract [1]. Although the specific etiology remain elusive, clinical and experimental data define CD as a complex trait [2] that arises from the interaction between the host's genetic background (susceptibility) [3] [4], mucosal immunity and the resident bacterial flora [5]. The most popular theories suggest that an imbalance in the immune system of people with CD leads to a local intolerance to the intestinal microbiota. During this process, it is the accumulation of white blood cells in the intestinal lining that leads to inflammation and ulceration.

Most ecological studies of CD have provided evidence of differences in the intestinal microbiota of healthy and diseased individuals [6-9]. However, despite intensive efforts, it is not known whether the bacteria associated with CD are primary etiologic agents of the disease, or secondary opportunistic colonizers that arise in a previously altered colon. Several lines of circumstantial evidence support an active role for bacteria in the development and/or progression of the disease: first, patients with inflammatory bowel disease (IBD) have increased amounts of bacteria attached to their gastrointestinal epithelial surfaces when compared to healthy individuals [10], and lesions in CD tend to appear in regions where bacterial counts are higher [11]: second, enteric bacteria can trigger the mucosal inflammation in IL-10, IL-2 and T cell receptor (TCR)-α deficient animal models [12], and none of these mutants developed intestinal disease when maintained under germ-free conditions: and third, specific agents such as *Mycobacterium avium* sbsp. *paratuberculosis* have been cultured from ulcers of CD [1, 13-18].

Key to understanding the role of bacteria in the etiology of CD is an accurate knowledge of the bacterial communities inhabiting such a complex ecosystem as the human colon. Our goal was to compare the microbiological profile of the intestinal mucosa from healthy volunteers and CD patients, and to identify any bacterial species characteristic of CD patients. DGGE is capable of separating nucleotide sequences that differ by as little as one nucleotide, and here we have used it to separate 16S rRNA gene fragments that correspond to the most dominant bacterial species present in the mucosal community. Sequence analysis of individual DGGE bands has also enabled us to identify individual bacterial species characteristic of CD patients.

Materials and Methods

Patients

Mucosal samples were obtained during colonoscopy from 19 CD patients with ileal, ileocolonic and colonic involvement, and 15 non-IBD controls (C). Thirty-six biopsies were taken from different locations of CD patients, 19 from inflammed and 17 from apparently normal mucosa, and 25 biospies from different locations from non-IBD controls (Table 7).

Table 7. Id codes of subjects and biopsies introduced in this study. Origin of the biopsies and mucosal state of the tissue are also indicated.

Patient	Age	Gender	Biopsy	Mucosa state	Biopsy location
C01	57	female	C01a	unaffected	colon
C02	53	female	C02a	unaffected	colon
C03	51	female	C03a	unaffected	colon
C04	36	female	C04a	unaffected	colon
C05	58	male	C05a	unaffected	colon
C06	29	female	C06a	unaffected	ileum
			C06b	unaffected	colon
C07	40	female	C07a	unaffected	colon
			C07b	unaffected	colon
			C07c	unaffected	colon
C08	51	female	C08a	unaffected	rectum
C09	53	male	C09a	unaffected	rectum
			C09b	unaffected	colon
C10	65	female	C10a	unaffected	colon
C11	48	female	C11a	unaffected	rectum
C12	50	female	C12a	unaffected	colon
			C12b	unaffected	rectum
			C12c	unaffected	colon
C13	46	male	C13a	unaffected	colon
C14	14	female	C14a	unaffected	colon
			C14b	unaffected	ileum
			C14c	unaffected	colon
			C14d	unaffected	rectum
C15	36	male	C15a	unaffected	colon
			C15b	unaffected	sigmoid colon
CD01	44	female	CD01a	unaffected	colon
			CD01b	affected	colon
CD02	26	female	CD02a	unaffected	ileum
			CD02b	affected	ileum
			CD02c	affected	ileum
CD03	27	male	CD03a	affected	colon
CD04	28	male	CD04a	unaffected	colon
			CD04b	affected	colon
CD05	32	female	CD05a	affected	colon
CD06	39	female	CD06a	unaffected	colon
			CD06b	affected	ileum
CD07	35	female	CD07a	unaffected	ileum
			CD07b	affected	colon
			CD07c	unaffected	colon
CD08	30	male	CD08a	unaffected	colon
CD09	35	male	CD09a	affected	duodenum
CD10	62	male	CD10a	unaffected	rectum
			CD10b	unaffected	ileum
			CD10c	affected	colon
CD11	48	female	CD11a	unaffected	colon
			CD11b	affected	colon
CD12	35	male	CD12a	unaffected	ileum
			CD12b	affected	ileum
CD13	23	female	CD13a	unaffected	colon
			CD13b	affected	ileum
CD14	25	female	CD14a	unaffected	colon
			CD14b	affected	colon
CD15	73	male	CD15a	affected	colon
			CD15b	unaffected	colon
CD16	15	male	CD16a	affected	ileum
			CD16b	unaffected	colon
CD17	44	female	CD17a	affected	not known
CD18	28	male	CD18a	affected	ileum
			CD18b	unaffected	colon
CD19	42	female	CD19a	unaffected	colon
			CD19b	affected	colon
IC01	49	female	IC01a	unaffected	not known
UC01	35	female	UC01a	unaffected	colon
			UC01b	unaffected	colon
UC02	27	male	UC02a	unaffected	ileum
			UC02b	unaffected	colon

Unaffected refers to non-inflammed tissue in non-IBD controls and apparently normal mucosa in IBD patients.
Affected refers to inflammed mucosa of IBD patients.

Control subjects underwent colonoscopy either for familial studies or hemorrhoidal rectorrhagia. On average, patients included in this study presented a 8.5 ± 2.2 year evolution, with CDAI, VanHees and CDEIS indexes of 203.5 ± 10.0, 123.8 ± 6.4, and 14.7 ± 3.1, respectively. Antibiotic treatment within the two months prior to colonoscopy was the only exclusion criterion. Samples were similarly obtained from two patients with UC (Ulcerative colitis) and one with IC (Ischemic colitis). Characteristics and clinical data of patients and non-IBD controls analyzed are shown in Table 8.

Table 8. Characteristics and clinical data of IBD patients and non-IBD controls analyzed.

Group	n	Age	Gender	Main characteristics	
C	15	45.8 ± 3.59	F: 73.3% M: 26.7%	familial CRC: constipation: hemorrhoids: UAD:	6.70% 26.70% 33.30% 33.30%
CD	19	36.7 ± 3.72	F: 52.6% M: 47.4%	years of disease: CDAI: VanHees: CDEIS: Medication:	8.5 ± 2.2 203.5 ± 10.0 123.8 ± 6.4 14.7 ± 3.1 52.63% Ø/local/mesalazine 5.26% corticosteroids 36.84% immunosuppressor 5.26% anti-TNF
UC	2	31 ± 2.37	F: 50%		
IC	1	27	Female		

C: Control; CD: Crohn's disease; UC: ulcerative colitis; IC: Ischemic colitis; UAD: Unspecific abdominal discomfort, CRC: Colorectal cancer

Bowel preparation and Sampling

Colon cleansing was ensured by using a biphosphate compound (Fleet Phospho Soda ®) and following the company's instructions. Two doses of 45 ml were given the day prior to the colonoscopy with the first dose taken in the late afternoon followed by 2 litres of water at little swallows. This procedure was repeated 4 hours later.

Biopsies were taken from ascendant, transversal and descendant colon, ileum, sigma and rectum. In ulcerative inflammations, biopsies were taken from both ulcerated and non-ulcerated mucosa. All samples were immediately placed in sterile tubes containing 0.5 ml of PBS and stored at –20ºC.

Sample treatment and DNA extraction

Biopsy samples were subjected to three mild ultrasounds-wash cycles in order to discard both transient and loosely attached bacteria. Each cycle consisted of 30 seconds at 50 Hz followed by a wash with 1 ml of PBS (phosphate buffered saline). After washing, tissue samples were then re-suspended in 0.5 ml of TENS buffer (0.05 M TrisHCl, 0.1 M EDTA pH=8, 0.1 M NaCl and 2% SDS) in the

presence of 1 ng μl^{-1} proteinase K and incubated overnight at 37ºC to remove cellular material. Afterwards, DNA was extracted and purified using the phenol-chloroform method [19, 20]. Two steps of Phenol:Chloroform:Isoamil acohol (25:24:1) and a further step of Chloroform:Isoamil acohol (24:1) were applied. DNA was precipitated with absolute ethanol and sodium acetate 3M. Pellets were resuspended in 10 mM TrisHCl, pH=7.4 and stored at –80 ºC until use. DNA concentration in all extracts was determined fluorimetrically by using PicoGreen™ (Molecular Probes, Eugene, Oreg.) and compared to a standard curve made up with DNA from herring sperm (Sigma-Aldrich).

Bacterial 16S rRNA gene amplification and Denaturing Gradient Gel Electrophoresis (DGGE)

For DGGE fingerprinting purposes, the 16S rRNA gene was partially amplified from extracted genomic DNA using universal bacterial primers GC-357F 5'-CGCCCGCCGCGCCCCGCGCCCGGCCCGCCGCCCCCGCCC-CCT ACG GGA GGC AGC AG-3' (341-357) [21] and 907R 5'-CCG TCA ATT CCT TTG AGT TT-3' (907-926) [22] flanking a ca. 580 bp fragment that includes variable regions V3 to V5. PCR reactions were prepared as follows: 5 µl 10 × Buffer (II) (Applied Biosystems, Foster City, CA), 3 µl $MgCl_2$ (25 mM; Applied Biosystems), 4 µl deoxyribonucleoside triphosphates (10 mM), 1 µl forward primer and 1µl reverse primer (25 pmol μl^{-1} each), 0.2 µl AmpliTaq® DNA polymerase (5 U μl^{-1}, Applied Biosystems) and 50 ng of genomic DNA as template in a total volume of 50 µl. PCR was performed using a GeneAmp® PCR System model 9700 cycler (Applied Biosystems). The following step-down cycling program was used: 2 min of initial denaturation at 96 ºC followed by 5 cycles of 30 sec at 94ºC, 1 min at 56 ºC (annealing) and 1min at 72 ºC plus 25 cycles of 30 sec at 94ºC, 1 min at 54ºC (annealing) and 1min at 72 ºC (extension) with a final extension of 10 min at 72 ºC. Amplification products were cleaned using Montage ™ PCR centrifugal filter columns (Millipore) and stored at -20 ºC. Products were visualized by gel electrophoresis on 1.5% (wt/vol) agarose gels in 0.5 × TBE buffer [45 mM Tris-borate (pH 8.0), 1 mM EDTA] and staining with ethidium bromide (0.5 µg ml^{-1}).

DGGE was carried out using an Ingeny phorU DGGE system (Vlissingen, The Netherlands) at 60 ºC in 0.5 × TAE buffer. Approximately 2 µg of DNA were loaded in each lane into a 6% polyacrylamide gel containing a vertical denaturing gradient ranging from 30% to 70% Urea/Formamide. The electrophoresis was run for 16 hours at 120V. The gel was then stained with ethidium bromide (1 µg ml^{-1}) for 30 min, placed on a UV transilluminator, and photographed. To standardize

the relative positions of bands, three lanes containing a mix of four 16S rRNA gene fragments from different known species were run in each DGGE.

Sequencing

After analyzing gel images, main bands were excised with a sterile scalpel. Pieces of polyacrylamide containing the band of choice were then placed on a tube with 200 µl of 10 mM Tris-HCl pH 7.4 and heated at 60ºC for 30 minutes and vigorously vortexed every 10 min. Supernatants were used for the reamplification of the 16S rRNA gene fragment contained in the excised band. PCR was performed as described above, but this time the forward primer 357F 5'- CCT ACG GGA GGC AGC AG-3' was used without the GC-clamp. PCR products were cleaned, and sequenced in both directions (forward and reverse) by Macrogen Inc (Seoul, Korea) with an ABI 3730XL Automatic DNA Sequencer.

Sequence editing and analysis

High-quality consensus sequences were obtained and manually refined using the Bioedit software package [23]. Alignments were carried out with ClustalW [24] software. Consensus sequences were compared to those in GenBank and the Ribosomal Database Project by using BLASTN 2.2.10 [25]. Sequences were grouped in Number of operational taxonomic units or phylotypes with the DOTUR program [26], using the furthest neighbor method at a precision level of 0.01, i.e. 99% minimum similarity for any pair of sequences to belong to the same phylotype, on a distance matrix with the Jukes-Cantor correction calculated with the DNADIST program of the Phylip software package [27].

GenBank accession numbers

Sequences obtained in this study were deposited in the EMBL/GenBank database under the accession numbers AM75626 through AM75766.

Statistics

Fingerprints on DGGE gels were translated into a presence-absence band matrix. Bands that fell at the same position in the gel and further confirmed to be the same sequence were considered 'genuine bands'. These bands were seacrhed in all patients and their presence or absence constructed the binary matrix. Statistical analysis was performed using the SPSSx version 11.0. Clustering of band patterns was performed using a hierarchical analysis by the Ward's [28] method based on the pattern difference coefficient. The same binary matrix was used in a correspondence analysis using those bands occurring two or more times. Significance of distances/similarities between groups was checked using an

ANOVA analysis. Pearson chi-square test was used to compare the prevalence of genus and species between controls and CD patients.

Results

DGGE fingerprinting comparison

The banding pattern obtained by DGGE represents the major constituents of the analyzed community [29]. Species that contribute less than 1% of the total population would not be readily detected by this molecular approach [21]. On a first visualisation, band patterns of samples from the same individual, either from affected or unaffected mucosa, were virtually identical. In addition the number of common, recognizable bands was found to be higher among the healthy individuals with some bands being consistent to most of the healthy individuals. By contrast, DGGE banding from CD patients, displayed uneven profiles, which were difficult to associate with a given pattern (Figure 12).

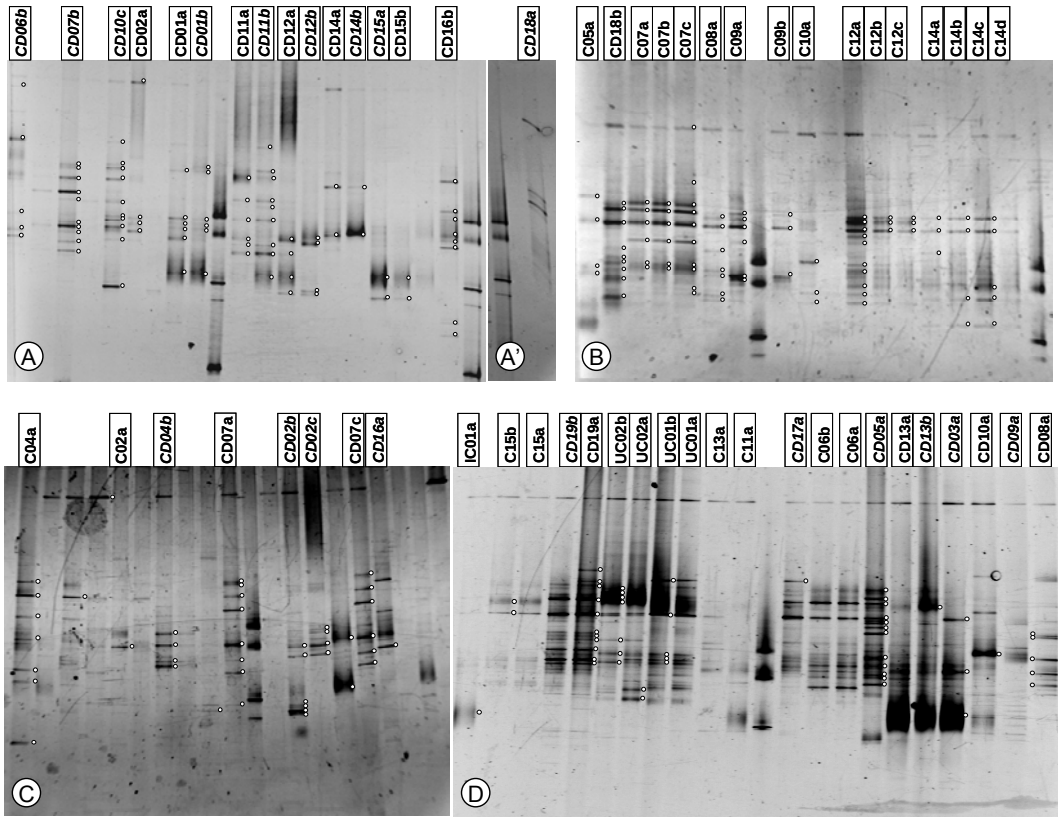


Figure 12. *Eubacteria* PCR-DGGE fingerprints of mucosa associated bacteria obtained from 15 non-IBD controls (C01-C15) and 19 CD patients (CD01-CD19). Different biopsies from the same patient are identified as '*a-d*' and biopsies obtained from inflammed mucosa of CD patients in italics. White dots indicate the main bands subjected to re-amplification and sequencing.

Hierarchical analysis of the banding patterns confirmed the initial visual observations; a first level of grouping was obtained after comparing DGGE profiles of the same individual (>99% similarity), regardless of the mucosal status (Figure 13). At the second level, two main clusters were obtained. In the first cluster, 64.7% of the profiles corresponded to samples from healthy individuals (84.6% of controls analyzed), 17.6% from CD patients, 11.7% from UC patients and 5.8% from IC patients. In the second cluster, 88.2% of the profiles corresponded to samples obtained from CD patients (83.3% of CD patients analyzed). Noticeably, of the three samples grouped in the first cluster, two corresponded to CD patients in remission.

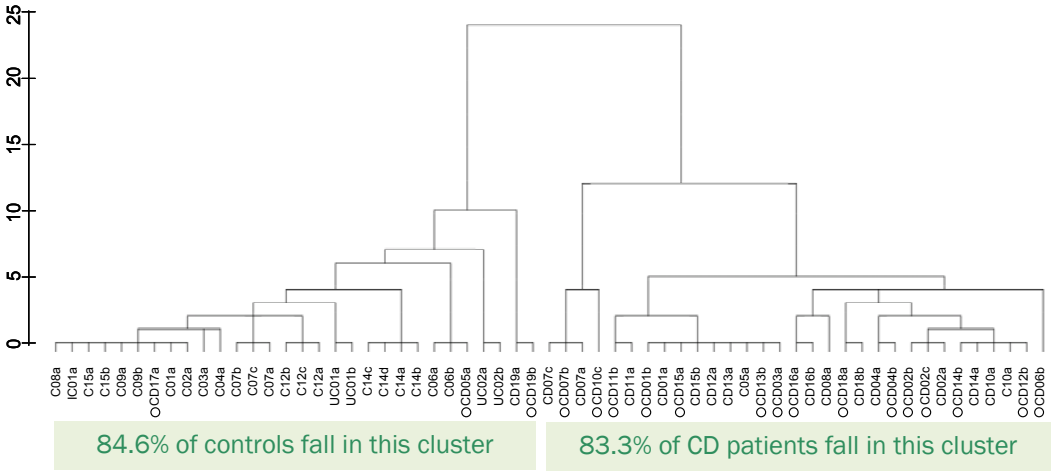


Figure 13. Hierarchical distance clustering of biopsy DGGE profiles. The scale bar describes DGGE similarity between profiles. C: non-IBD control; CD: Crohn's Disease; UC: ulcerative colitis; IC: Ischemic colitis; ◯: samples from CD lesions (inflammed mucosa); *a-d*: different biopsies from the same patient.

The existence of two clusters, with a healthy pattern clearly differentiated from the CD one, was statistically corroborated by a correspondence analysis (Figure 14, $p<0.001$). In addition, a larger inter-patient variability ($P<0.001$) within the CD cluster was observed, which confirms the observation that DGGE profiles from CD were more heterogeneous and disperse (Figure 15). This variability was not correlated with any patient variable such as age, gender, smoking habit, disease activity index, age at the onset of disease, years of evolution and treatment.

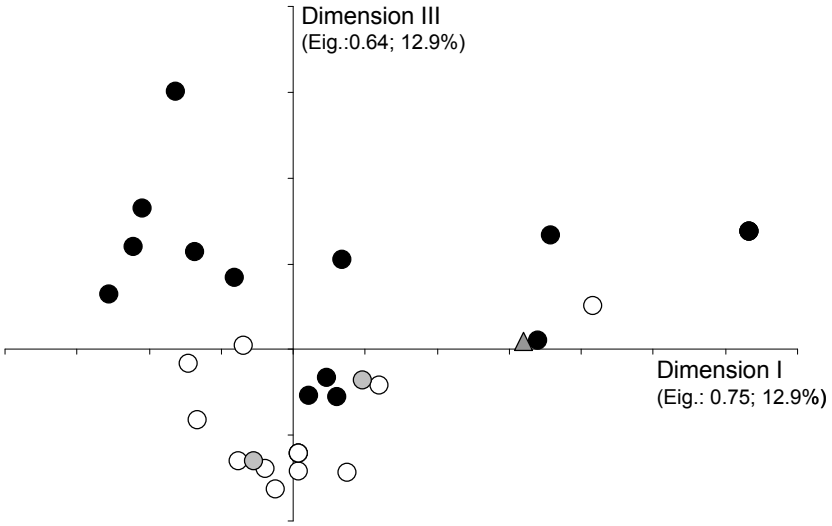


Figure 14. Correspondence analysis using DGGE profiles. ◯: non-IBD controls; ●: Crohn's disease (CD) patients; ●: Ulcerative colitis (UC); ▲: ischemic colitis (IC). Dimension 3 separates specimens according to their diagnostic (P <0.001) and dimension 1 displays the higher DGGE profile dispersion within CD patients.

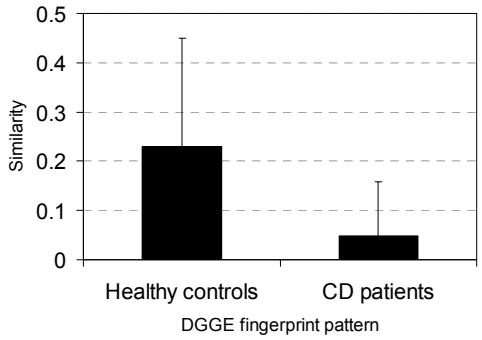


Figure 15. Internal DGGE pattern consistency of non-IBD controls and CD specimens, using Dice similarity indexes.

Sequence analysis and bacterial composition

One hundred forty-one partial 16S rRNA gene sequences were obtained from both non-IBD and CD samples. The results of their best BLAST matches can be consulted in annexed Table 1. These fell into 58 different phylotypes (≤ 99% sequence identity from one phylotype to the next), of which 8 were novel (sharing ≤ 95% similarity with any GenBank sequence) (Table 9). *Firmicutes* (69%) and *Bacteroidetes* (13.8 %) were the dominant phylogroups, as expected from other studies of the gut microbiota [30-32]. By contrast, *Proteobacteria*, *Actinobacteria* and *Fusobacteria* were remarkably less abundant accounting for 6.9%, 5.1%, and 5.1% of phylotypes retrieved, respectively.

The genus *Faecalibacterium* was found to be significantly more prevalent in non-IBD controls than in CD patients (P=0.029) (Table 10). In particular,

Faecalibacterium prausnitzii (formerly *Fusobacterium prausnitzii* [33]), was found in 12 out of 15 healthy subjects (80%) whereas the prevalence in CD samples was significantly reduced to 42.1%. As for particular phylotypes, AM075691 was 98% similar to *Faecalibacterium prausnitzii*, and was found in 80% of controls and only in 15.8% of CD patients (P<0.001). Other sequences belonging to the same phylotype as AM075670, whose closest related bacterium also is *Faecalibacterium prausnitzii*, were more frequently found in the healthy cluster (P< 0.05).

Table 10. Prevalence of bacterial taxonomic groups in CD patients (N=19) and controls (N=15).

Genus	C	CD	*P*	Family	C	CD	*P*	Phylum	C	CD	*P*
Atopobium	1	0	0.253	*Coriobactericeae*	1	1	0.863	*Actinobacteria*	1	2	0.694
Collisinella	0	1	0.367								
Corynebacterium	0	1	0.367	*Corynebacteriaceae*	0	1	0.367				
Bacteroides	3	4	0.940	*Bacteroidaceae*	3	4	0.940	*Bacteroidetes*	5	6	0.914
Alitispes	1	2	0.694	*Rikenellaceae*	2	2	0.801				
Rikinella	1	0	0.253								
Clostridium	**1**	**8**	**0.020**	*Clostridiaceae*	13	13	0.203	*Firmicutes*	13	16	0.616
Faecalibacterium	**12**	**8**	**0.029**								
Enterococcus	0	2	0.195	*Enterococcaceae*	0	2	0.195				
Eubacterium	4	5	0.982	*Eubacteriaceae*	4	5	0.982				
Catenibacterium	1	1	0.863	*Lachnospiraceae*	7	12	0.336				
Roseburia	0	1	0.367								
Ruminococcus	6	11	0.300								
Staphylococcus	0	1	0.367	*Staphylococcaceae*	0	1	0.367				
Fusobacteriunm	1	2	0.591	*Fusobacteriaceae*	1	2	0.591	*Fusobacteria*	1	2	0.591
Enterobacter	0	1	0.367	***Enterobacteriaceae***	**1**	**7**	**0.039**	***Proteobacteria***	**1**	**8**	**0.020**
Escherichia	**1**	**6**	**0.074**								
Klebsiella	0	1	0.367								
Proteus	0	1	0.367								
Haemophilus	0	1	0.367	*Pasteurellaceae*	0	1	0.367				

Of the *Proteobacteria*, all those identified belong to the gamma subgroup and mainly to *Enterobacteriaceae* family, were found more consistently in CD patients (P = 0.020 and P = 0.039 respectively). At the genus level, *Clostridium* (P = 0.020) and *Escherichia* (P < 0.075) were found more frequently in CD patients. Taking into consideration the high heterogeneity of *Clostridium* genus, the prevalence of this organism was calculated according to their phylogenetic affiliation [34] and their pathogenic behavior. The most relevant difference corresponded to Cluster XIV a, which was more prevalent in CD patients (P = 0.042). This cluster is known to contain the less harmful species. Nevertheless, pathogenic species belonging to this cluster have been recently reported [35]. The sequences of some opportunistic pathogenic γ-*Proteobacteria* such as *Enterobacter*, *Proteus*, *Haemophilus* and *Klebsiella* were occasionally found in

CD patients, but never in the healthy subjects. As for the phylotypes, sequences corresponding to the same phylotype as AM075639 sequence (98% similarity to *Ruminococcus torques*) were only found in CD patients (P = 0.059).

On BLAST analysis against the GenBank, 25.5% of the 141 sequences showed a match with previously cultivated bacteria, whereas the other 74.5% were similar to environmental sequences from bacteria that have never been cultivated. Of these a vast majority were sequences obtained from human (88.7%) or animal (7.6%) fecal or intestinal microbiota. A minor portion of sequences (3.7%) was similar to anaerobes from other environments. Notably, some of the sequences (CD0302, CD1103 and CD1701) were similar to those also retrieved from CD stools in a previous study [7].

Discussion

We used a PCR-DGGE-sequencing approach to show that the profile of dominant bacteria present in the intestinal mucosa of CD patients differs from that of healthy subjects. Moreover, within the group of CD patients the patient-to-patient variability is greater than that seen within the group of healthy subjects (Figure 14). The etiology of CD is characterized in part by the genetic makeup of the individual, and mutations in several genes, including *NOD2* and other *IBD* genes, predispose the carrier to imbalances in the mucosal immune system leading to CD [2, 36, 37]. We postulate that individuals that are predisposed to CD are less able to regulate the microbial makeup of their intestines and this leads to an unstable microbial population, reflected in the large variability seen within the group of CD patients. This is borne up by the observation that the microbial profiles vary much less between one healthy individual and the next, implying a host factor in maintaining the microbial population of healthy individuals. To what extent the microbial imbalance contributes to, or reflects the damaging inflammatory response seen in CD patients is not clear. A good candidate for such host factors might be for example the alpha-defensins that constitute the mucosal immune system, and described recently [38]. Recent reports have implied a role for bacterial components such as flagellins, lipopolysaccharides, peptidoglycan or lipoproteins as part of the causal chain of events leading to the disease symptoms [39] and therefore a specific role for one or more bacteria in the ethiology of this disease cannot be ruled out.

Table 9. Different phylotypes found as calculated by DOTUR with the furthest neighbour method using a 99% similarity threshold.

DGGE band	Phylum	OTUs*	Accession number †	Similarity (%)	Nearest sequence	Description	Source
CD0605	*Bacteroidetes*	1	AM075633	94‡	AY986341	Uncultured bacterium clone D741	human intestinal microbiota
CD0404	*Firmicutes*	5	AM075643	94‡	AY916138	Uncultured bacterium clone KS90	human intestinal microbiota
CD1801	*Firmicutes*	1	AM075677	94‡	AY452007	Uncultured bacterium clone Muc3-18	human intestinal microbiota
CD1802	*Firmicutes*	1	AM075676	95‡	AY916258	Uncultured bacterium clone E108	human intestinal microbiota
CD1908	*Firmicutes*	3	AM075733	95‡	AF371609	Uncultured bacterium clone p-334-a3	animal intestinal microbiota
UC0203	*Bacteroidetes*	1	AM075736	95‡	AY895203	*Bacteroides thetaiotaomicron* strain 8736	Pure culture
CD0505	*Firmicutes*	1	AM075746	95‡	AF132260	Uncultured bacterium adhufec310	human intestinal microbiota
CD0801	*Firmicutes*	1	AM075756	95‡	AY977866.1	Uncultured bacterium clone LF64	human intestinal microbiota
CD0704	*Firmicutes*	1	AM075631	96	AY981791	Uncultured bacterium isolate HuCC43	human intestinal microbiota
C0704	*Bacteroidetes*	1	AM075669	96	AY126616	*Bacteroides massiliensis* strain B84634	Blood cluture of newborn
C0404	*Bacteroidetes*	1	AM075694	96	AY919925	Uncultured bacterium clone Eldhufec050	human intestinal microbiota
CD0406	*Firmicutes*	1	AM075700	96	AF153858	Uncultured bacterium adhufec80.25	human intestinal microbiota
CD0211	*Firmicutes*	1	AM075714	96	AJ315487	Uncultured bacterium isolate HuCC43	human intestinal microbiota
UC0104	*Firmicutes*	2	AM075740	96	AY452007	Uncultured bacterium clone Muc3-18	human intestinal microbiota
CD1603	*Firmicutes*	1	AM075637	97	DQ057466	Uncultured bacterium ic1337	animal intestinal microbiota
C0304	*Firmicutes*	1	AM075688	97	AY471654	Uncultured bacterium clone Adhufec015khh	Crohn's disease patients microbiota
CD0716	*Firmicutes*	2	AM075706	97	AY916138	Uncultured bacterium clone KS90	human intestinal microbiota
CD0206	*Firmicutes*	2	AM075708	97	X94964	*Ruminococcus schinkii* strain Bie 41	Rumen of suckling lambs
CD0507	*Bacteroidetes*	1	AM075743	97	AY986226	Uncultured bacterium clone D593	human intestinal microbiota
CD0502	*Firmicutes*	1	AM075748	97	X94967	*Ruminococcus gnavus*	bacterium from the rumen
CD0601	*Firmicutes*	4	AM075626	98	AY920077	Uncultured bacterium clone Eldhufec202	human intestinal microbiota
CD0714	*Firmicutes*	4	AM075651	98	AY983968	Uncultured bacterium clone MC49	human intestinal microbiota
CD0103	*Firmicutes*	2	AM075655	98	AY850358	*Enterococcus faecalis* strain SFL 16S	human intestinal microbiota
C0302	*Firmicutes*	3	AM075679	98	AY984355	Uncultured bacterium clone B256	human intestinal microbiota
C0305	*Firmicutes*	1	AM075687	98	AF530354	Uncultured bacterium clone cadhufec18d05sav	Crohn's disease patients microbiota
CD1401	*Firmicutes*	5	AM075632	99	AY986349	Uncultured bacterium clone D750	human intestinal microbiota
CD0705	*Firmicutes*	6	AM075634	99	AY985177	Uncultured bacterium clone C012	human intestinal microbiota
CD1102	*Fusobacteria*	1	AM075635	99	M58686	*Fusobacterium varium* ATTC 8501	Type strain
CD1601	*Firmicutes*	2	AM075636	99	AY305313	Butyrate-producing bacterium SM7/11	human intestinal microbiota
CD0101	*Firmicutes*	8	AM075639	99	AY452007	Uncultured bacterium clone Muc3-18	human intestinal microbiota
CD1003	*Firmicutes*	1	AM075641	99	Y12669	*Clostridium perfringens*	Pure culture
CD0405	*Firmicutes*	6	AM075644	99	AF153858	Uncultured bacterium adhufec80.25	human intestinal microbiota

Table 9. (*Continued*)

CD0408	*Firmicutes*	4	AM075645	99	AJ315487	Uncultured bacterium isolate HuCC43	human intestinal microbiota
CD0102	*Proteobacteria*	9	AM075657	99	AF527825	*Escherichia coli* strain RREC III	Bovine feces
CD1103	*Firmicutes*	1	AM075659	99	AF530331	Uncultured bacterium clone cadhufec20a04yvb	Crohn's disease patients microbiota
CD1402	*Proteobacteria*	1	AM075660	99	AY362908	*Haemophilus parainfluenzae* strain CCUG 12836	Pure culture
CD1201	*Proteobacteria*	1	AM075662	99	AJ301682	*Proteus mirabilis* strain CIP1031181T	Pure culture
C0502	*Firmicutes*	1	AM075668	99	AY983727	Uncultured bacterium clone NQ77	human intestinal microbiota
C1001	*Fusobacteria*	1	AM075675	99	AY684430	Uncultured bacterium clone HuRC28	human intestinal microbiota
C0402	*Firmicutes*	2	AM075691	99	AY986207	Uncultured bacterium clone D569	human intestinal microbiota
CD0702	*Bacteroidetes*	2	AM075702	99	AY985751	Uncultured bacterium clone C783	human intestinal microbiota
CD0504	*Bacteroidetes*	1	AM075742	99	AY643082	*Alistipes finegoldii* strain 4401054	Blood culture, patients with colon cancer
CD1101	*Bacteroidetes*	4	AM075630	100	AY985581	Uncultured bacterium clone C568	human intestinal microbiota
CD1604	*Firmicutes*	7	AM075638	100	X94967	*Ruminococcus gnavus* ATCC29149	Type strain
CD0607	*Firmicutes*	1	AM075647	100	AY452003	Uncultured bacterium clone Muc3-10	human intestinal microbiota
CD0201	*Firmicutes*	1	AM075653	100	AY305314	Butyrate-producing bacterium SM4/1	human intestinal microbiota
CD0202	*Firmicutes*	2	AM075654	100	Y10028	*Clostridium* sp. strain DR6A	animal intestinal microbiota
CD1005	*Proteobacteria*	1	AM075656	100	AY292865	*Klebsiella pneumoniae* isolate 521	Human feces
CD0717	*Actinobacteria*	1	AM075658	100	AY388411	Uncultured bacterium clone Adhuright30	human intestinal microbiota
C0702	*Firmicutes*	9	AM075670	100	AY985486	Uncultured bacterium clone C447	human intestinal microbiota
C0902	*Firmicutes*	3	AM075678	100	AY592220	Uncultured bacterium Kazan-4B-06/BC19-4B-06	Kazan mud volcano
C0405	*Actinobacteria*	1	AM075695	100	AJ251324	*Atopobium oviles*	Pure culture
CD0207	*Actinobacteria*	4	AM075710	100	AJ012838	*Corynebacterium simulans* strain UCL557	Human clinical samples
CD1904	*Firmicutes*	2	AM075732	100	AY985152	Uncultured bacterium clone BB76	human intestinal microbiota
CD1007	*Firmicutes*	1	AM075752	100	AY030342.1	*Staphylococcus epidermidis* strain KL-096 16S	Pure culture
CD0802	*Fusobacteria*	1	AM075754	100	AY684430.2	Uncultured bacterium clone HuRC28	human intestinal microbiota
UC0206	*Firmicutes*	1	AM075761	100	AY985642.1	Uncultured bacterium clone C643	human intestinal microbiota
CD1701	*Firmicutes*	1	AM075765	100	AF530337	Uncultured bacterium clone cadhufec17c02sav	Crohn's disease patients microbiota

‡ Novel phylogroups with identities ≤95% to any NCBI/EMBL sequence are indicated. BLAST performed in 2006.
***** Number of sequences of this particular phylotype
† Representative for the phylotype

Although the CD patients analyzed were being treated with different medications; mesalazine (9), corticoids (1), moderate immunosuppressors (7), and anti-TNFs (1), no effect of medication on the microbiota composition can be deduced from our data. We also confirm previous findings that the bacterial composition does not change significantly along the tract of the large intestine [40], irrespective of the disease state of the individual.

Sequence analysis revealed that some species have different prevalence in the CD and in the healthy clusters. In addition to *Faecalibacterium prausnitzii*, a butyrate-forming bacterium [41] which was found to be characteristic of non-IBD controls, the higher presence of γ-*Proteobacteria* in CD samples also help to explain the observed hierarchical grouping obtained from DGGE profile analysis. In particular, *Enterobacteriaceae* were more prevalent in agreement with previous works [6-8]. *Clostridium* spp. higher prevalence in CD patients can be a result of an opportunistic development of this genus in an altered colonic ecosystem. Some other opportunistic pathogens such as *Proteus*, *Haemophilus*, and *Enterobacter* have been found only in the group consisting of CD patients. Of particular note, *Escherichia coli* was found in 31.6% of CD patients, compared with just 6.7 % in healthy people (P = 0.074). This finding corroborates recent and early research suggesting a possible link between this bacterium and the pathogenesis of Crohn's Disease [42-49], and it has been reported that infection by *E. coli* O157:H7 may mimic right colonic Crohn's disease in its presentation [50].

Interestingly, the four biopsies of mucosa obtained from the two UC patients fall into the group of what we considered a "healthy" pattern. A posterior analysis of the eubacterial-DGGE fingerprints of 5 additional UC patients and 5 CD patients by using the software GelComparII was performed in other to corroborate this result (complete data not shown in this section). Using these data two main clusters appeared separating UC patients from CD patients with the exception of one CD patient that grouped in the "UC" cluster (Figure 16). Although more data are needed, this implies that CD and UC might actually be similar clinical manifestations of different problems. In agreement with this, recent studies have shown that patients with CD differ from those with UC in their immunological response to bacterial flagellins [51], supporting our tentative observation that the microbiological manifestation of these diseases appear to differ from each other. One might speculate that, in IBD, the colon microbiota is not an epiphenomenon of the inflammatory process.

Molecular studies on the bacterial composition of colonic mucosa from CD patients using fresh biopsies are scarce. Most similar works published to date used either stools or fixed histological material from Pathological Anatomy laboratories as specimens. An important methodological issue of this work was to ensure that bacteria to be analyzed were those intimately associated with the mucosa, and that transient luminal bacteria were removed prior to any analysis. This obeys to the assumption that mucosa-associated bacteria are more likely to be involved in the inflammatory response of the colon than luminal counterparts, due to their close proximity to the host epithelium. Moreover, any bacterium should be found intimately associated to the mucosa to be considered a causative agent of CD and it should be in a number high enough to be detected by molecular methods.

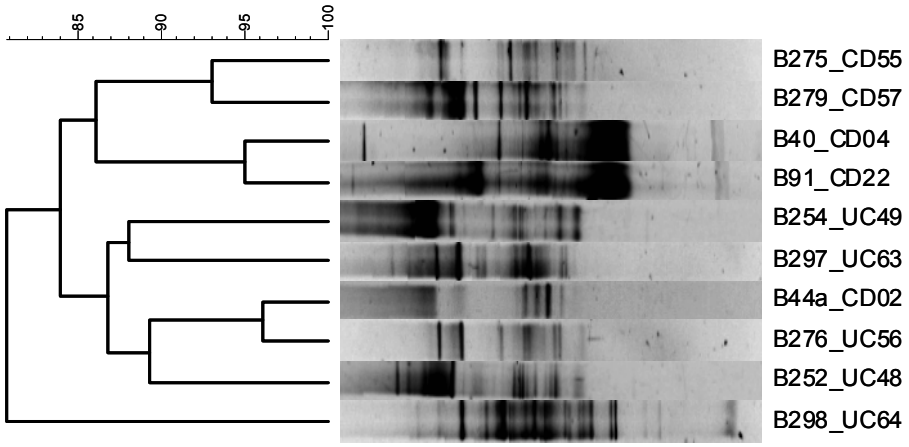


Figure 16. Hierarchical cluster of Eubacterial-DGGE fingerprints of 5 CD patients and 5 UC patients by using GelComparII.
Similarity coefficient: Different bands; Dendogram type: Ward; Optimisation: 1.00%; Tolerance: 0.5%

Concluding remarks and clinical implications

Our results indicate that CD patients harbor a specific microbiota, which is distinguishable from that of healthy subjects. A deeper knowledge of the CD microbiology may help in the search for an adequate therapeutic strategy. Moreover, the different microbiota in CD mucosa may indicate a possible relationship between some bacteria and the etiology of this disease.

Finally, the different prevalence of *Escherichia coli*, *Clostridium* and *Ruminococcus torques* could potentially serve as indicators of CD. In contrast *Faecalibacterium prausnitzii* can be considered as a good indicator of healthy colonic mucosa. Thus, we found that the simultaneous presence of *Clostridium* spp. and *Escherichia coli* and absence of *Faecalibacterium* are as much as 100

times more likely to be found in CD patients than in healthy people. With further studies, it might also be possible to use these techniques to differentiate CD from UC and might be useful as complementary information to help with diagnosis of the inflammatory bowel diseases.

References

1. Shanahan F, O'Mahony J. The mycobacteria story in Crohn's disease. Am J Gastroenterol. 2005; 100:1537-1538

2. Hugot JP, Chamaillard M, Zouali H, et al. Association of NOD2 leucine-rich repeat variants with susceptibility to Crohn's disease. Nature. 2001;411:599-603

3. Cho JH, Nicolae DL, Gold LH, et al. Identification of novel susceptibility loci for inflammatory bowel disease on chromosomes 1p, 3q, and 4q: Evidence for epistasis between 1p and IBD1. PNAS. 1998;95:7502-7507

4. Hugot JP, Alberti C, Berrebi D, et al. Crohn's disease: the cold chain hypothesis. Lancet. 2003;362:2012-2015

5. Elson CO. Genes, microbes, and T cells--new therapeutic targets in Crohn's disease. N Engl J Med. 2002;346:614-616

6. Seksik P, Rigottier-Gois L, Gramet G, et al. Alterations of the dominant faecal bacterial groups in patients with Crohn's disease of the colon. Gut. 2003;52:237-242

7. Mangin I, Bonnet R, Seksik P, et al. Molecular inventory of faecal microflora in patients with Crohn's disease. FEMS Microbiol Ecol. 2004;50:25-36

8. Manichanh C, Rigottier-Gois L, Bonnaud E, et al. Reduced diversity of faecal microbiota in Crohn's disease revealed by a metagenomic approach. Gut. 2006; 55:205-211

9. Prindiville T, Cantrell M, Wilson KH. Ribosomal DNA sequence analysis of mucosa-associated bacteria in Crohn's disease. Inflamm Bowel Dis. 2004;10:824-833

10. Swidsinski A, Ladhoff A, Pernthaler A, et al. Mucosal flora in inflammatory bowel disease. Gastroenterol. 2002;122:44-54

11. Sartor RB. Postoperative recurrence of Crohn's disease: the enemy is within the fecal stream. Gastroenterol. 1998;114:398-400

12. Rennick DM, Fort MM. Lessons from genetically engineered animal models. XII. IL-10-deficient (IL-10(-/-) mice and intestinal inflammation. Am J Physiol Gastrointest Liver Physiol. 2000; 278:829-833

13. Bull TJ, McMinn EJ, Sidi-Boumedine K, et al. Detection and verification of *Mycobacterium avium* subsp. *paratuberculosis* in fresh ileocolonic mucosal biopsy specimens from individuals with and without Crohn's disease. J Clin Microbiol. 2003;41:2915-2923

14. Roholl PJ, Herrewegh A, van Soolingen D. Positive IS900 in situ hybridization signals as evidence for role of *Mycobacterium avium* subsp. *paratuberculosis* in etiology of Crohn's disease. J Clin Microbiol. 2002;40:3112

15. Greenstein RJ. Is Crohn's disease caused by a mycobacterium? Comparisons with leprosy, tuberculosis, and Johne's disease. Lancet Infect Dis. 2003;3:507-514

16. Ryan P, Bennett MW, Aarons S, et al. PCR detection of *Mycobacterium paratuberculosis* in Crohn's disease granulomas isolated by laser capture microdissection. Gut. 2002;51:665-670

17. Ryan P, Kelly RG, Lee G, et al. Bacterial DNA within granulomas of patients with Crohn's disease-detection by laser capture microdissection and PCR. Am J Gastroenterol. 2004;99:1539-1543

18. Van Kruiningen HJ. On the use of antibiotics in Crohn's disease. J Clin Gastroenterol. 1995;20:310-316

19. Moore DD. Purification and concentration of DNA from aqueous solutions. New York: John Wiley & Sons. Inc; 1996

20. Sambrook J, Fritsch EF, Maniatis T. Molecular cloning: a laboratory manual. 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY: 1989

21. Muyzer G, de Waal EC, Uitterlinden AG. Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA. Appl Environ Microbiol. 1993;59:695-700

22. Muyzer G, Teske A, Wirsen CO, et al. Phylogenetic relationships of *Thiomicrospira* species and their identification in deep-sea hydrothermal vent samples by denaturing gradient gel electrophoresis of 16S rDNA fragments. Arch Microbiol. 1995;164:165-172

23. Hall T. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucl Acids Symp Ser. 1999;41:95-98

24. Thompson JD, Gibson TJ, Plewniak F et al. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res.1997;25:4876-4882

25. Altschul SF, Madden TL, Schaffer AA, et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 1997;25:3389-3402

26. Schloss PD, Handelsman J. Introducing DOTUR, a computer program for defining operational taxonomic units and estimating species richness. Appl Environ Microbiol. 2005;71:1501-1506

27. Felsenstein J. PHYLIP (Phylogeny Inference Package) version 3.5c. Distributed by the author. Department of Genetics, University of Washington, Seattle; 1993

28. Jobson JD editor. Applied multivariate data analysis. Vol 2. New York, NY: Springer-Verlag; 1992

29. Heuer H, Smalla K. Application of denaturing gradient gel electrophoresis (DGGE) and temperature gradient gel electrophoresis (TGGE) for studying soil microbial communities. In: van Elsas JD, Wellington EMH, and Trevors JT, ed. Modern soil microbiology. New York, NY: Marcel Dekker, Inc; 1997:353-373

30. Hold GL, Pryde SE, Russell VJ, et al. Assessment of microbial diversity in human colonic samples by 16S rDNA sequence analysis. FEMS Microbiol Ecol. 2002;39:33-39

31. Wang X, Heazlewood SP, Krause DO, et al. Molecular characterization of the microbial species that colonize human ileal and colonic mucosa by using 16S rDNA sequence analysis. J Appl Microbiol. 2003;95:508-520

32. Eckburg PB, Bik EM, Bernstein CN, et al. Diversity of the Human Intestinal Microbial Flora. Science. 2005;308:1635-1638

33. Duncan SH, Hold GL, Harmsen HJM, et al. Growth requirements and fermentation products of *Fusobacterium prausnitzii*, and a proposal to reclassify it as *Faecalibacterium prausnitzii* gen. nov., comb. nov. Int J Syst Evol Microbiol. 2002;52:2141-2146

34. Collins MD, Lawson PA, Willems A, et al. The phylogeny of the genus *Clostridium*: proposal of five new genera and eleven new species combinations. Int J of Syst Bacteriol. 1994;44:812-826

35. Finegold SM, Song Y, Liu C, et al. *Clostridium clostridioforme*: a mixture of three clinically important species. Eur J Clin Microbiol Infect Dis. 2005;24:319-324

36. Ogura Y, Bonen DK, Inohara N, et al. A frameshift mutation in NOD2 associated with susceptibility to Crohn's disease. Nature. 2001;411:603-606

37. McGovern D, Ahmad T. New IBD genes? Gut. 2005;54:1060-1061

38. Eckmann L. Defence molecules in intestinal innate immunity against bacterial infections. Curr Opin Gastroenterol. 2005;21:147-151

39. Akira S, Takeda K. Toll-like receptor signalling. Nat.Rev.Immunol. 2004;4:499-511

40. Zoetendal EG, von Wright A, Vilpponen-Salmela T, et al. Mucosa-Associated Bacteria in the Human Gastrointestinal Tract Are Uniformly Distributed along the Colon and Differ from the Community Recovered from Feces. Appl Environ Microbiol. 2002;68:3401-3407.

41. Hold GL, Schwiertz A, Aminov RI, et al. Oligonucleotide Probes That Detect Quantitatively Significant Groups of Butyrate-Producing Bacteria in Human Feces. Appl Environ Microbiol. 2003;69:4320-4324

42. Martin HM, Campbell BJ, Hart CA, et al. Enhanced *Escherichia coli* adherence and invasion in Crohn's disease and colon cancer. Gastroenterol. 2004;127:80-93.

43. Darfeuille-Michaud A, Boudeau J, Bulois P, et al. High prevalence of adherent-invasive *Escherichia coli* associated with ileal mucosa in Crohn's disease. Gastroenterol. 2004;127:412-421.

44. Keighley MR, Arabi Y, Dimock F, et al. Influence of inflammatory bowel disease on intestinal microflora. Gut .1978;19:1099-1104.

45. El-Tawil AM. High incidence of Crohn's disease in pre-menopausal women: *Escherichia coli* may be contributing. Int J Colorectal Dis. 2004;19:498

46. Darfeuille-Michaud A. Adherent-invasive Escherichia coli: a putative new *E. coli* pathotype associated with Crohn's disease. Int J Med Microbiol. 2002;292:185-193

47. Glasser AL, Boudeau J, Barnich N, et al. Adherent invasive *Escherichia coli* strains from patients with Crohn's disease survive and replicate within macrophages without inducing host cell death. Infect Immun. 2001;69:5529-5537

48. Boudeau J, Glasser AL, Masseret E, et al. Invasive ability of an *Escherichia coli* strain isolated from the ileal mucosa of a patient with Crohn's disease. Infect Immun. 1999;67:4499-4509

49. Darfeuille-Michaud A, Neut C, Barnich N, et al. Presence of adherent *Escherichia coli* strains in ileal mucosa of patients with Crohn's disease. Gastroenterol. 1998;115:1405-1413

50. Ilnyckyj A, Greenberg H, Bernstein CN. *Escherichia coli* O157:H7 infection mimicking Crohn's disease. Gastroenterol. 1997;112:995-999

51. Lodes MJ, Cong Y, Elson CO, et al. Bacterial flagellin is a dominant antigen in Crohn disease. J Clin Invest. 2004;113:1296-1306

Chapter 2

Characterisation of *Escherichia coli* populations associated to the intestinal mucosa of Crohn's disease patients and control subjects

Introduction

The etiopathogenesis of inflammatory bowel diseases (IBDs) remains uncertain; however, several genetic, immunologic, and environmental factors have been implicated [1- 5]. Among environmental factors, microorganisms have been extensively reported to be involved in the onset or perpetuation of inflammation [6-11]. Intestinal microbiota may be involved in Crohn's disease (CD) in two ways: i) a low-grade infection by a persistent pathogen, either traditional or opportunistic; ii) dysbiosis of the commensal microbiota, in which protective bacteria decrease as harmful bacteria increase.

Several studies on fecal [12, 13] and mucosa-associated bacterial communities [14-16] have shown that the microbiota of patients with CD differ from those of healthy controls, as well as those of patients with ulcerative colitis (UC) [17, 18]. Although the reported changes are not always consistent, numbers of *Proteobacteria*, in particular *Escherichia coli*, are generally increased, whereas *Firmicutes* are scarcer in CD patients [12-14, 18, 19]. Whether these changes are causative factors [20] or consequences of inflammation [21] remains controversial.

To date, several pathogens have been proposed as causative agents. In particular, adherent-invasive *E. coli* (AIEC) is achieving increasing relevance since it has been reported to be more prevalent in CD patients than in controls in several countries: France [22], the United Kingdom [23], and the USA [19, 24]. Moreover, new pathogenic traits have been recently discovered [25-27].

In this work, we compared the *E. coli* subtype diversity in the human gut of healthy individuals to that of CD patients by analyzing the clonality of approximately 100 *E. coli* isolates per patient. All subtypes were further characterized by focusing on genetic and phenotypical pathogenic features. Furthermore, all *E. coli* isolates were tested to assess whether they belonged to the AIEC group, thus obtaining the prevalence, abundance, and richness of this pathovar. Finally, the collection of AIEC strains was seropathotyped.

Materials and Methods

Clinical data of patients and sample treatment

Patients with clinically confirmed IBD were recruited between 2002 and 2007 in two hospitals. The subjects were not exposed to antibiotics for 2 months prior to

colonoscopy. Three sets of IBD patients and controls were subjected to three methodological approaches (Table 11). Control subjects were asymptomatic and did not present inflammation and/or evidence of polyps during colonoscopy. Rectorrhagia or hemorrhoids (60%), irritable bowel syndrome (28%), diverticulitis (3%), and colorectal cancer patients' relatives (9%) comprised the control group. Among CD patients, 39% had Crohn's colitis (C-CD), 35% had Crohn's ileitis (I-CD), and 26% had ileal/colonic disease (IC-CD). Patients with different levels of endoscopic activity (from remission to severe inflammation) and on different medications for the treatment of IBD were included.

Biopsies were taken from the ileum and/or colon with sterile forceps, immediately placed in sterile tubes without any buffer, and maintained at 4ºC for *E. coli* isolation and at –20ºC for DNA extraction. Biopsies were subjected to three mild ultrasound-wash cycles in order to discard both transient and loosely attached bacteria. Each cycle consisted of 30 seconds at 50 Hz, followed by washing with 1 ml of 1× PBS (phosphate buffered saline).

Table 11. Sample size, clinical data and biopsy origin of subjects.

			Gender	Zone sampled (N patients/ *N biopsies*)			
Diagnose	N	Age (Mean±SD)	(% Males)	ileum	colon	ileum + colon	duodenum
Set used for *E. coli* quantification by RTi-PCR							
C	17	46.4 ± 13.6	23%	1 / *1*	14 / *21*	2 / *4*	
CD	26	36.4 ± 12.8	48%	4 / *7*	16 / *23*	5 / *11*	1 / *2*
UC	8	31.2 ± 11.1	62.5%	–	7 / *8*	1 / *2*	
Set used for analysis of diversity and characterisation of the entire *E. coli* population							
C	12	49 ± 20	42%	2 / *2*	10 / *10*	–	
CD	10	34 ± 8.5	30%	1 / *1*	7 / *7*	2 / *4*	
Set used for analysis of diversity and characterisation of the AIEC population							
C	28	44.7 ± 16.1	43%	9 / *9*	11 / *11*	8 / *16*	
CD	20	33.5 ± 9.4	30%	4 / *4*	9 / *9*	7 / *14*	

E. coli quantification by Real-Time PCR

DNA extraction and quantification was performed as described previously [14].

The previously reported 16S rDNA-based primers and probe were used for *E. coli* quantification [28]. The PCR conditions were slightly modified as follows: amplification reactions were carried out in a total volume of 20 µl containing 1× Taqman Buffer A, 6 mM $MgCl_2$, 200 µM dNTPs, 300 nM of each primer, 100 nM of probe, and 1 U of AmpliTaqGold (Applied Biosystems). Treatment of AmpliTaqGold with DNAase I was performed to avoid contamination with residual *E. coli* DNA. DNAase I was used at 5×10^{-2} U/µl and incubated at room temperature for 5 minutes, followed by a deactivation step carried out by heating

at 99ºC for 10 minutes. Amplification of PCR reactions was carried out in a 7500 Real-Time PCR System (Applied Biosystems).

Samples were analyzed in triplicate, and an additional tube with an inoculum of DNA of known concentration was added as internal standard in order to detect the possibility of reaction inhibition. Ct values were translated into colony-forming units (CFU). To normalize the biopsy size, human cell quantification was performed using the control RTi-CKFT-18S kit (Eurogentec).

Isolation of mucosa-associated *E. coli* from fresh biopsies

Biopsies were directly streaked onto Tryptone Bile X-Glucuronide Medium (TBX, Oxoid). In order to release any intracellular bacteria, a mild osmotic shock was applied to biopsies by incubation for 5 minutes in distilled water, during which time eukaryotic cells are disrupted, while bacterial cells resist the osmotic shock (data not shown). Afterwards, the biopsies and a fraction of the supernatants were also cultured in TBX. All colonies were collected and confirmed using the indole assay.

Molecular characterisation of isolates:

Rep-PCR: The initial step in subtyping

The first step in analyzing *E. coli* clonality was performed using the Insertion Element IS3-Based PCR subtyping method and primers IS3A and IS3B together, as described previously [29]. Rep-PCR profiles were compared using the GelComparII software (Applied Maths). The clonality of one isolate being representative of each cluster was further confirmed by PFGE. An example of the dendrogram resulting from the IS3-based Rep-PCR profiles of *E. coli* isolates from a single patient is shown in annexed Figure 17.

Pulsed-field gel electrophoresis: confirmation of clonality and molecular epidemiology

PFGE was performed as described elsewhere [30] using two different enzymes: *Xba*I and *Spe*I. Restriction reactions were incubated at 37ºC for at least 4 h, according to the manufacturer's instructions (Takara Bio Inc.). Electrophoresis was carried out in a CHEF-DR® III System (Bio-Rad). TIFF images were analyzed using GelComparII software (Applied Maths). Similarity indices were estimated using the Dice method, with a band position tolerance of 1.5%. A pairwise distance matrix among subtypes was the basis for cluster analysis by UPGMA. The similarity percentage cut-off to distinguish clonally distinct groups

was 90%. A clone type defined for each group of clones was further characterized by phylotyping and seropathotyping.

Phylotyping and virulence genotyping by PCR

Determination of the major *E. coli* phylogenetic group (A, B1, B2 and D) was performed as previously described by Clermont *et al.* [31].

Virulence genotyping of all *E. coli* subtypes was performed as described elsewhere [32, 33] using primers specific for 11 genes that encode extraintestinal virulence factors characteristic of ExPEC. These included six adhesins (pyelonephritis-associated pili (*papC*), S and F1C fimbriae (*sfa/focDE*), afimbrial Dr-binding adhesins (*afa/draBC*), type 1 fimbriae (*fimH*), and type 1 variant of avian pathogenic *E. coli* strain MT78 ($fimAv_{MT78}$)); three toxins (*hlyA*, *cnf1*, and *cdtB*); and one aerobactin gene (*iucD*). They also included two protectin/invasion-encoding genes that corresponded to K1 *kps* variant (*neuC*) and brain microvascular endothelial cell invasion gene (*ibeA*). Specific genes for diarrhoeagenic *E. coli* pathovars were also screened (*stx1*, *stx2*, *eae*, *bfpA*, *ipaH*, *pCDV432*, *eltA,* and *est*).

Phenotypic characterisation:

Serotyping

Determination of O and H antigens was carried out using the method previously described by Guinée *et al.* [34].

Adhesion and invasion assays in Intestine-407 epithelial cells

Cell culture of the epithelial cell model Intestine-407 (ATCC CCL-6™) was performed as described previously [35].

A first **qualitative screening** of the invasion capacity of all isolates was performed in 96-well plate cell cultures. Intestine-407 was seeded at a density of 1×10^5 cells/well and incubated for 20 hours. Before infection, cell monolayers were washed twice with 100 µl of PBS and 100 µl of EMEM medium (Cambrex) supplemented with 10% heat-deactivated FBS (Cambrex) was added. Inoculation of bacteria was performed with a 96-deep-well replica plater (Edge BioSystems). Three hours after infection, the medium was replaced with fresh medium containing 100 µg/ml gentamicin (Sigma) and incubated for 1 hour. The cells were then lysed with 100 µl of 100% TritonX-100 (Sigma). Five-microliter spots of direct cell lysates were applied to a square LB plate. Once grown, the spots were classified in four categories according to density, from 0 to 3, ranging from less to more invasive (Figure 17). A portion of isolates of densities 2 and 3 was

quantitatively confirmed (50% of category 3 isolates with a maximum of five per patient and 10% of category 2). Some isolates with a density of 1 were also checked, but none were confirmed to be invasive.

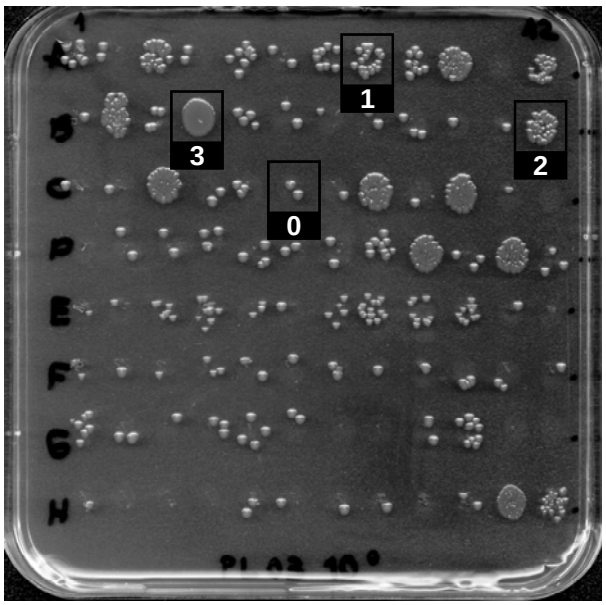


Figure 17. Qualitative invasion assay in 96-well plates. Four categories of invasion capacity were defined accordingly to density of the spots: 0, 1, 2, and 3, from less to more invasive.

Quantitative analyses of adhesion and invasion properties were performed in 24-well plates as described previously [22]. Briefly, 24-well plates containing 4×10^5 cells/well incubated for 20 hours were infected at a multiplicity of infection of 10. Duplicated plates, for adhesion and invasion assays were incubated for 3 hours at 37ºC. For bacterial adhesion assays, cell monolayers were washed 5 times with PBS and lysed with 1% Triton X-100. Adhered bacteria were quantified by plating them in nutrient agar. Plating was performed in a maximum period of 30 minutes to avoid bacterial lysis by Triton X-100. Adherence ability (I_ADH) was determined as the mean number of bacteria per cell. For bacterial invasion assays, monolayers were washed twice with PBS after 3 hours of infection, and fresh cell culture medium containing 100 μg ml^{-1} of gentamicin was added for 1 hour to kill extracellular bacteria. After cell lysis with 1% Triton X-100, the number of intracellular bacteria was also determined by plating. All assays were performed in triplicate. The invasive ability was expressed as the percentage of intracellular *E. coli* compared with the initial inoculum, taken as 100%: I_INV (%) = (intracellular bacteria / 4×10^6 bacteria inoculated) × 100.

Survival and replication in J774 macrophages

The macrophage-like J774A.1 cell line (ATCC TIB-67™) was used as model for *E. coli* survival and replication assays. Adherent and invasive *E. coli* isolates were evaluated for their ability to survive and replicate inside macrophages, as

previously described [36]. Macrophages were seeded at 2×10^5 cells per well in two 24-well plates and incubated for 20 hours. Once overnight medium was removed and fresh medium was added, bacteria were seeded at a multiplicity of infection of 10. Centrifugation at 900 rpm for 10 minutes, plus an additional incubation at 37ºC for 10 minutes, was performed to assist the internalisation of bacteria within macrophages. Phagocytosed bacteria were killed with gentamicin (20 μg ml^{-1}), and intracellular bacteria were quantified as for invasion assays after 1 and 24 hours of infection. All assays were performed in triplicate. Results were expressed as the mean percentage of the number of bacteria recovered after 1 and 24 h post-infection compared with the initial inoculum, taken as 100%: I_REPL (%) = (cfu ml^{-1} at 24h / cfu ml^{-1} at 1h)× 100. Those strains with I_INV > 0.1 and I_REPL > 100% were classified as AIEC in this study.

Statistical analyses

Quantitative parameters, such as richness, abundance and diversity, were compared by one-way ANOVA. In those cases for which the interaction between several factors was of interest, a factorial ANOVA was applied. Diversity was calculated using the Shannon Index (H'). For prevalence values, we used the Pearson-X^2 test. Data with high variability between subjects were log transformed prior to statistical analysis.

Results

Mucosa-associated *E. coli* in the human gut: abundance, subtype richness and diversity, and characterisation of strains

The subset of patients included for *E. coli* quantification is summarized in Table 11. Mucosa-associated *E. coli* numbers were significantly higher in CD patients than in controls and UC patients, as revealed by quantitative PCR (Figure 18A). In particular, patients with Crohn's ileitis harboured higher counts in comparison with C-CD and IC-CD patients, regardless of the zone sampled (Figure 18B). No differences were observed between the four categories defined inside the control group. In UC patients, when the activity of colitis and the onset of the disease were contrasted, 10 times less *E. coli* abundance was detected on newly diagnosed patients (P=0.02), and no correlation with the activity was observed.

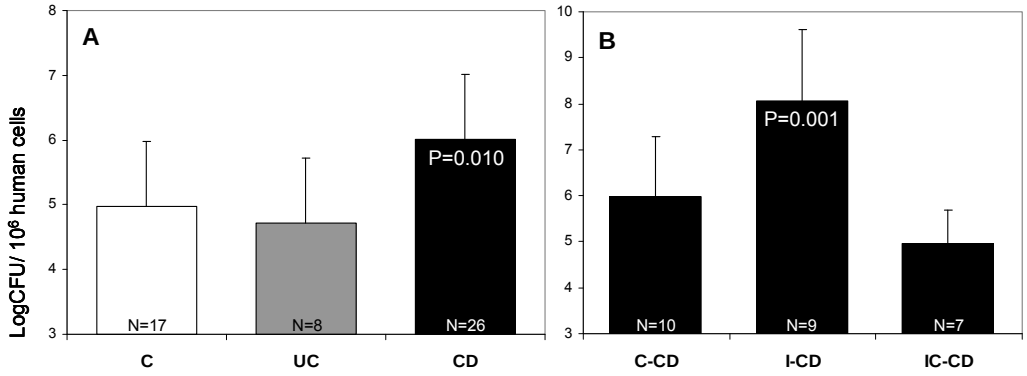
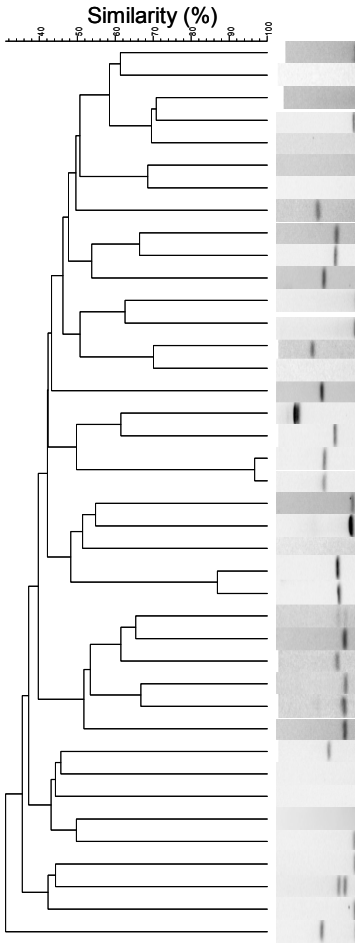


Figure 18. Mucosa-associated *E. coli* quantification by RTi-PCR. A: On average, *E. coli* was approximately 10 times more abundant in Crohn's disease (CD) patients than in controls (C). B: The subgroup Crohn's ileitis (I-CD) showed higher quantities of *E. coli* than other subgroups within CD patients.

To characterize the *E. coli* population, the clonality of 1,769 isolates obtained from 12 controls and 10 CD patients (Table 11) was analyzed, resulting in a total of 40 *E. coli* subtypes. No genetic relatedness was observed between *E. coli* subtypes from CD patients or healthy subjects, as evidenced by macro-restriction analysis. Figure 19 shows the pulsotypes* of the 40 ECGsubtypes and their relative abundance within the *E. coli* population of a given patient, serotype, phylogenetic origin, AIEC determination, and ExPEC-like virulence gene carriage. A mean value of 2.0 ± 1.0 *E. coli* subtypes (subtype richness) was found in the mucosa of healthy controls. Similarly, CD patients harboured from 1 to 6 subtypes (2.3 ± 1.5). In terms of diversity, which takes into account the relative abundance of subtypes (see annexed Table 2), both groups of subjects also showed similar characteristics (CD: H' = 3.1 ± 3.1; C: H' = 4.5 ± 5.3). The majority of *E. coli* subtypes were found to be unique to any given patient; only 4 were simultaneously detected in two individuals of the same group (C or CD).

Typically, *E. coli* subtypes were consistently found in the intestinal mucosa sampled along the intestinal tract, in both ulcerated and non-ulcerated tissue.

No differences in virulence gene frequency were found between *E. coli* subtypes isolated from CD patients (CD-*EC*) and healthy controls (C-*EC*) (Table 12), except for a decreased prevalence of the *iucD* gene in CD-*EC* (P=0.027). The majority of mucosa-associated *E. coli* subtypes carried at least one adhesin (95% and 91% in C-*EC* and in CD-*EC*, respectively), with *fimH* being the most common (95% and 90%, respectively), followed by *papC* (55% and 40%, respectively). Around 30% of *E. coli* subtypes carried at least one gene encoding a toxin (28% in C-*EC* and 34% in CD-*EC*).



ECGtype	Patient/s	Abundance (%)	Presence of AIEC isolates	Serotype	Phylogroup	ExPEC-like
ECG-021	CD109	47 (88.7)	–	O54:H21	B1	–
ECG-028	CD111	160 (100)	+	ONT	D	+
ECG-034	CD109	6 (11.3)	–	O127:H28	D	–
ECG-042	CD79	21 (91.3)	–	O7:H6	B2	–
ECG-036	C125	1 (1.5)	+	O6:HNT	B2	+
ECG-013	C127	2 (2.2)	–	O18:H1	B2	+
ECG-012	C71	1 (1.2)	–	O18:H1	B2	+
ECG-063	CD110	5 (4.7)	–	ONT:H21	B1	–
ECG-008	C124	1 (1.1)	–	O14:H7	B2	+
ECG-046	C81	11 (100)	–	ONT:H29	B1	–
ECG-023	C119	74 (82.2)	+	ONT:H–	Atypical	–
ECG-016	C82	10 (90.9)	–	O55:H–	A1	+
ECG-017	C83	26 (100)	–	O22:H18	B2	–
ECG-031	C119	24 (17.8)	–	O161:H4	D	–
ECG-049	C84	94 (100)	–	O15:H16	B2	+
ECG-004	C124 / C125	77 (88.5) / 64 (98.5)	+	O22:H7	B1	–
ECG-019	CD79	2 (8.7)	–	ONT:H–	A	–
ECG-018	CD55	34 (100)	–	ONT:H–	A	–
ECG-025	C70 / C71	21 (48.8) / 9 (10.6)	–	O6:H31	B2	+
ECG-024	C71	33 (38.8)	–	O174:H–	B1	–
ECG-026	CD112	168 (100)	–	O1:H–	B2	+
ECG-037	C82	1 (9.1)	–	O1:H–	D	+
ECG-015	CD113	84 (100)	–	O2:H6	B2	+
ECG-001	CD53	160 (90.9)	+	O6:H31	B2	+
ECG-002	CD53	1 (0.6)	–	O175:H–	B1	–
ECG-057	CD110	9 (8.5)	–	O11:H18	D	+
ECG-054	C124	9 (10.3)	–	O14:H–	D	+
ECG-053	CD110	7 (6.6)	+	O11:H–	D	+
ECG-056	CD110	80 (75.5)	–	ONT:H18	D	+
ECG-059	CD110	4 (3.8)	–	ONT:H18	D	+
ECG-055	C127	88 (97.8)	–	O17:H18	D	+
ECG-047	CD61	1 (0.9)	–	O119:H21	D	–
ECG-009	CD53 / CD61	15 (8.5) / 1 (0.9)	–	O83:H1	B2	+
ECG-041	C69	95 (100)	–	O2:H–	B2	+
ECG-043	C80	116 (100)	–	O83:H1	B2	+
ECG-005	CD72	25 (65.8)	–	O4:H5	B2	+
ECG-022	C70	1 (2.3)	–	O15:H–	A	+
ECG-065	CD61	112 (98.2)	–	O141:H–	A	–
ECG-060	C70 / C71	21 (48.8) / 42 (49.4)	–	O102:H6	D	+
ECG-064	CD72	13 (34.2)	–	O166:H21	B1	–

Figure 19. Consensus UPGMA dendrogram generated from Dice coefficients of *SpeI* and *XbaI* PFGE profiles of the 40 *E. coli* clonal types detected (ECGtypes). Listed are their abundances within the *E. coli* population for each patient, as well as their genetic and phenotypic characterisation.

The ExPEC-like column indicates the presence of more than two virulence genes characteristic of ExPEC. Nd: not determined. Nt: Non typeable.

Table 12. Summary of the genetic and phenotypic characterisation of *E. coli* subtypes isolated from 12 healthy subjects (C) and 10 Crohn's disease (CD) patients.

Characteristic		*E. coli* subtypes isolated from 12 C (n=20)	*E. coli* subtypes isolated from 10 CD (n=20)	P
EXPEC-like strains		70%	50%	NS
Virulence genes richness		3.4 ± 1.7	2.9 ± 1.9	NS
Adhesins	*papC*	55%	40%	NS
	sfa/focDE	20%	15%	NS
	afa/draBC	5%	15%	NS
	fimH	95%	90%	NS
	$fimAv_{MT78}$	25%	20%	NS
Capsule formation	*neuC*	15%	20%	NS
Iron transport	*iucD*	75%	40%	0.027
Invasion	*ibeA*	10%	15%	NS
Toxins	*hlyA*	20%	15%	NS
	cnf1	20%	15%	NS
	cdt	5%	5%	NS
DEC	EAEC	5%	0%	NS
Phylogroup	A	10.5%	15%	NS
	B1	15.8%	20%	NS
	B2	42.1%	30%	NS
	D	31.6%	35%	NS

The prevalence of virulence genes and the phylogroup origin expressed as percentages. The mean ± standard deviation is provided for general virulence gene carriage. ExPEC-like: strains with ≥2 virulence genes typical from extraintestinal pathogenic *E. coli*. NS: no significant differences.

Conversely, several genes were characteristic of certain phylogroups. In particular, *E. coli* subtypes belonging to the B2 phylogroup carried a higher richness of virulence genes (B2-*EC*: 4.7 ± 0.3; D-*EC*: 3.20 ± 0.3; A-*EC*: 2.0 ± 0.5; B1-*EC*: 1.0 ± 0.5; P = 0.000). A high prevalence of ExPEC-like subtypes (more than 2 virulence genes characteristic of ExPEC) was also characteristic of B2- and D-*EC* subtypes (87% of B2-*EC*; 75% of D-*EC*; 40% of A-*EC*; 0% of B1-*EC*; P = 0.001). However, the phylogroup distribution was similar between CD patients and controls, with B2 and D being the most prevalent in both groups. Consequently, the virulence gene content of mucosa-associated *E. coli* was similar between CD patients and controls.

Comparison of AIEC prevalence and subtype abundance, richness, and diversity between CD patients and controls

The presence of AIEC was screened in over 4,314 *E. coli* isolates obtained from ileum and colon samples of 20 CD patients and 28 controls (Table 11). Approximately 353 colonies were identified as putative invasive *E. coli* by the qualitatively analysis of invasion (categories 2 and 3), of which 136 were selected for confirmation as AIEC by quantitative methods (Table 13). The 53 isolates finally confirmed as AIEC were used for posterior analysis of clonality by PFGE.

The prevalence of AIEC was higher in CD patients than in controls, both in the colon and ileum (Figure 20), regardless of the localisation of the disease.

The abundance of AIEC, defined as the percentage of AIEC within the *E. coli* population, was also higher in CD patients than in controls (P = 0.05). On average, for colonic samples, AIEC isolates represented 5.34% of the *E. coli* population in CD patients and 0.95% in controls; however, variability was notable in both cases (SD values of 12.27% and 2.68%, respectively). In the ileum, the difference in abundance was not as great as that in the colon between CD patients and controls (CD: 4.84% ± 13.53; C: 3.58% ± 12.51) (Figure 20). Among CD patients, an increased abundance of AIEC was observed especially in I-CD, although it was not statistically significant.

A total of 22 AIEC subtypes were identified by PFGE (Figure 21). AIEC richness (number of AIEC subtypes per patient) was higher in CD patients than in controls (CD: 0.8 ± 1.4; C: 0.2 ± 0.4; P = 0.015) (Figure 20). In contrast, due to the large number of cases presenting only one subtype per patient, thus with a Shannon diversity index of zero, no significant differences between CD and C subjects were found in terms of AIEC subtype diversity.

Within a given patient, the same AIEC subtypes were detected in ulcerated and non-ulcerated mucosa (Figure 22). Furthermore, in most cases, several isolates with an identical pulsotype* demonstrated different adherence and invasive properties (Figure 23).

AIEC and non-AIEC characterisation

The seropathotypes* and phylogroups of 16 CD-AIEC subtypes and 6 C-AIEC subtypes obtained in the present study are shown in Table 14. Two serogroups, O6 and O22, comprised approximately 45% of the AIEC subtypes. However, a high variability of O:H serotypes and pulsotypes* was observed within the AIEC population.

Seventy-three percent of the AIEC subtypes were classified as ExPEC-like and none of them carried the signature virulence genes for the six diarrhoeagenic *E. coli* pathotypes. We compared the phylo-pathotype of the 22 AIEC subtypes with a subset of 38 mucosa-associated *E. coli* without adherent-invasive properties derived from the previous study of the entire *E. coli* population (non-AIEC). Interestingly, the frequency and distribution of virulence genes was similar between AIEC subtypes isolated from CD patients and controls, as well as between AIEC subtypes and non-AIEC subtypes. The B2 phylogroup was more prevalent within the AIEC population in comparison with non-AIEC subtypes (P = 0.044) (Table 15).

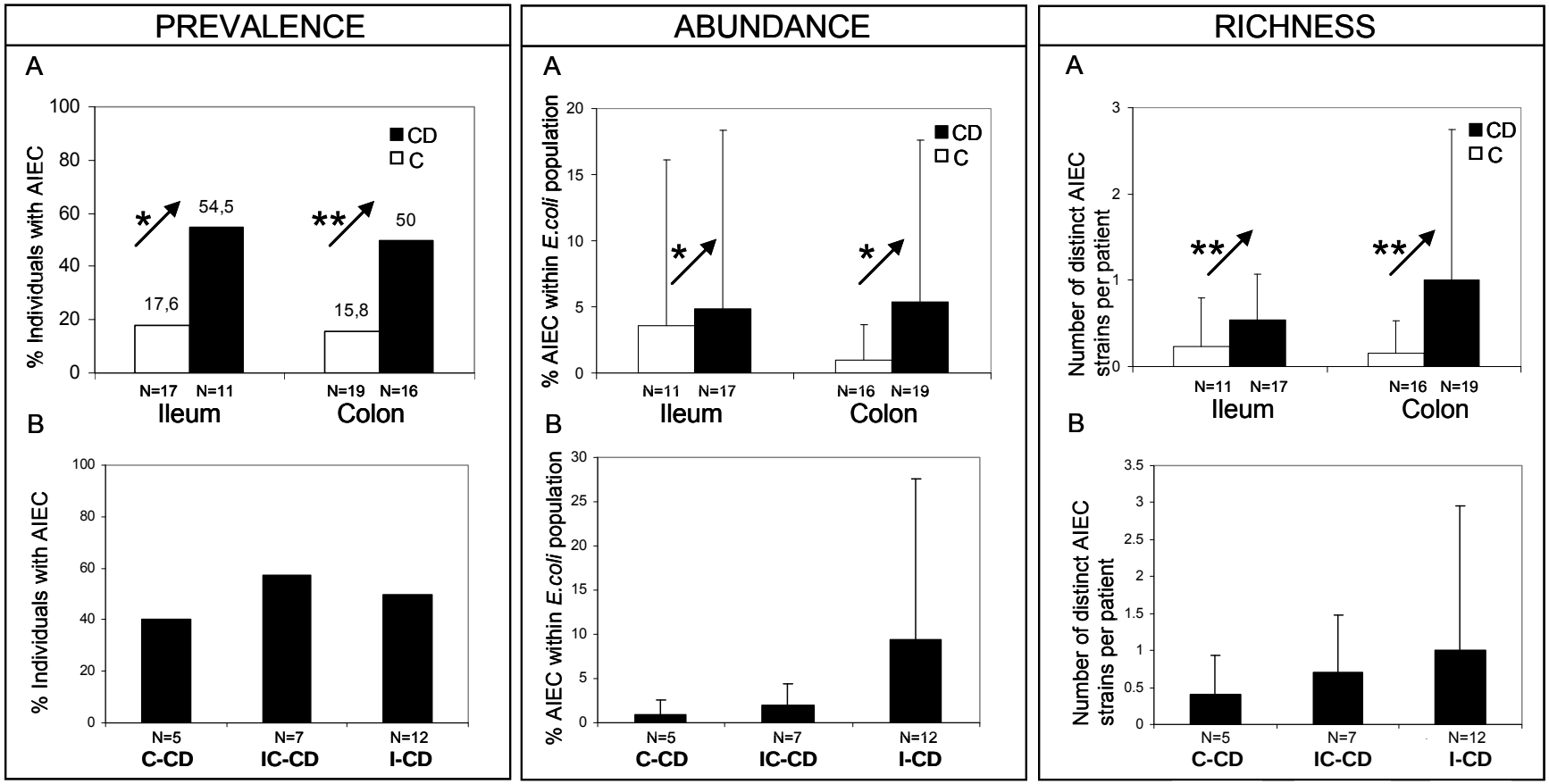
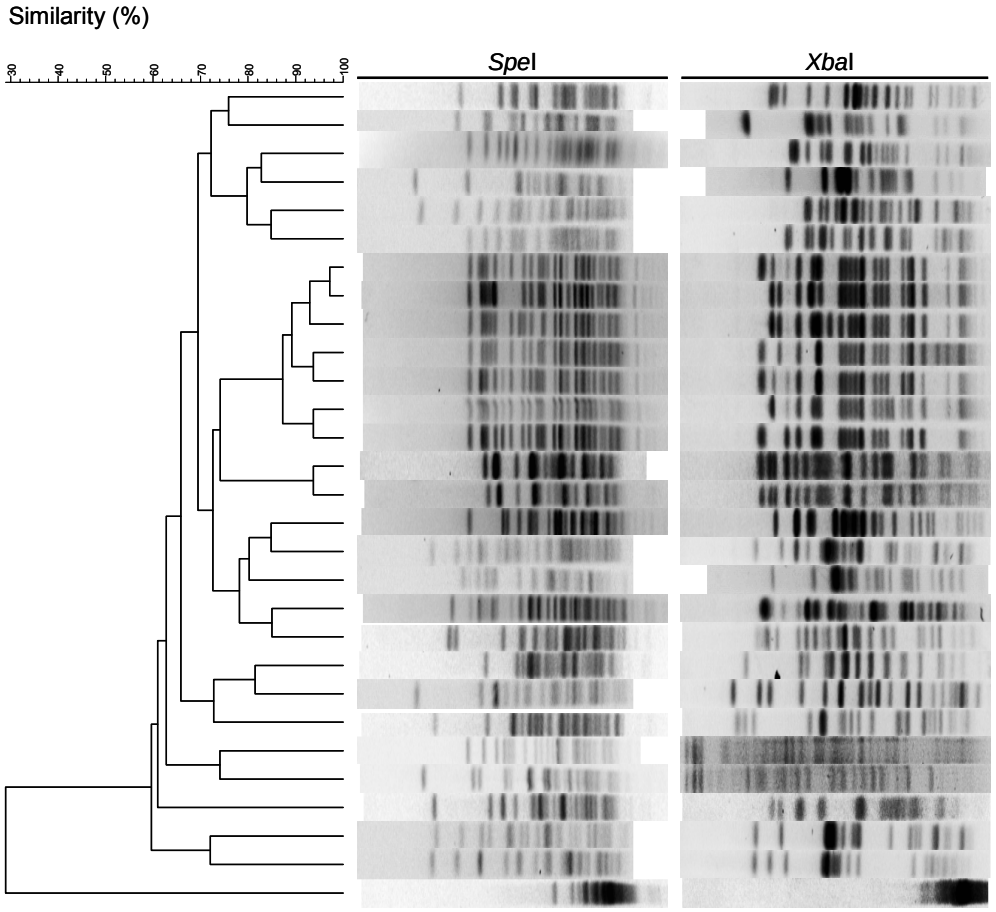


Figure 20. AIEC prevalence (% of patients with AIEC), AIEC abundance (% AIEC isolates /Total *E. coli* isolates) and AIEC richness (Number of different AIEC subtypes per patient) of 20 Crohn's disease (CD) patients and 28 healthy controls (C). B: AIEC prevalence, abundance and richness of 9 patients with Crohn's ileitis (I-CD), 6 with ileo-colonic disease (IC-CD) and 5 with Crohn's colitis (C-CD). Statistical significance is denoted as: * P = 0.05; ** 0.05 > P > 0.001.

Table 13. Estimation of adherent-invasive *E. coli* abundance in the ileum and colon of CD patients and controls.

Diagnose	Localisation	N	Id patient	[a] Frequency of putative invasive *E. coli*	[b] Freq. of confirmed invasive *E. coli*	[c] Estimated abundance of invasive *E. coli*	Range invasion indices	Range replication indices	[d] Estimated abundance of AIEC
CD	Ileum	4	120	2/91	2/2/2	2/91	0.284	1567 and 1688	2/91
			123	4/91	1/4/4	1/91	0.568	2362	1/91
			79	0/97					
			150	5/92	0/4/5	0/92			
	Ileum + colon	7	109	i: 1/60 - c: 1/66	0/2/2				
			110	i: 6/91 - c: 0/39	2/3/6	i:4/91	0.125 to 0.284	343 to 592	4/91
			111	i: 1/91 - c:5 /80	i:1/1/1 - c:4/4/5	i:1/91 - c:5/80	i: 0.625 - c:0.211 to 2.636	1053 to 2174	i:1/91 - c:5/80
			112	i: 0/88 - c: 0/91					
			113	i: 1/63 - c: 2/87	0/3/3				
			114	i: 1/91 - c: 9/86	i:1/1/1 - c:4/6/9	i:1/91 - c:6/86	i:0.117 - c:0.159 to 0.802	704 to 2733	i:1/91 - c:6/86
			132	i:41/90 - c:50/88	i:7/7/41 - c:1/7/50	i:41/90 - c:7/88	i:0.2 to 1.4 - c:0.148	660 to 981	i:41/90 - c:7/88
	Colon	9	61	3/148	1/3/3	1/148	0.259	150	1/148
			72	0/40					
			77	2/2	1/2/2	1/2	0.109	1297	1/2
			89	2/24	1/2/2	1/24	0.142	100	1/24
			92	0/92					
			118	11/81	0/11/11				
			122	2/31	2/2/2	2/31	0.216 and 0.309	1625 and 2562	2/31
			115	0/8					
			53	14/273	12/12/14	14/273	0.170 to 2.627	639 to 1726	14/273
C	Ileum	9	124	2/93	2/2/2	2/93	0.565 and 0.663	1692 and 2296	2/93
			125	47/91	5/5/47	47/91	0.285 to 3.330	293 to 1561	47/91
			126	1/14	1/1/1	1/14	0.226	1413	1/14
			156	2/92	0/2/2	0/92			
			158	0/23					
			159	1/92	0/1/1	0/92			
			163	2/70	0/2/2	0/70			
			164	3/92	0/3/3	0/92			
			165	0/30					
	Ileum + colon	8	128	i: 2/7 - c: 6/28	i:0/2/2 - c:3/5/6	c:3/28	c:0.134 to 0.177	2332 to 2749	c:3/28
			117	i: 0/20 - c: 0/22					
			130	i:0/16 - c:0/9					
			142	i: 16/91 - c: 9/91	i: 0/4/16 - c: 1/4/9	c:2/91	c:0.172	100	2/91
			144	i:2/69 - c:3/82	i:0/2/2 - c:0/2/3	i:0/69 - c:0/82			
			151	i:0/92 - c:0/62					
			152	i:2/90 - c:7/92	i:0/2/2 - c:0/5/7	i:0/90 - c:0/92			
			133	i:0/18 - c:0/13					
	Colon	11	69	0/73					
			71	0/96					
			70	0/96					
			81	0/14					
			80	0/156					
			82	0/16					
			83	0/96					
			84	0/96					
			119	66/91	1/14/66	4.7/91	0.111	1568	4.7/91
			127	17/91	0/2/17				
			88	2/94	0/2/2				

[a] Number of isolates of category 3 and 2/ Total *E. coli* isolates. i: ileum; c: colon. [b] Num. of confirmed invasive *E. coli*/ Num. of putative invasive *E. coli* tested/ Total putative invasive *E. coli*.
[c] Num. of confirmed invasive *E. coli*/ Total *E. coli* isolates. [d] Num. of confirmed AIEC/ Total *E. coli* isolates.



AIEC type	Patient	Sample origin	AIEC isolates belonging to the same clone*
AIEC05	CD114	COLON / ILEUM	5,5 / 1
AIEC06	C128	COLON	3
AIEC04	C125	ILEUM	43,24
AIEC07	C124 / C125	ILEUM / ILEUM	2 / 3,76
AIEC02	CD114	COLON	1
AIEC01	CD120	ILEUM	2
AIEC14-1			
AIEC14-2	CD132	COLON / ILEUM	5,85 / 11,7
AIEC14-3			
AIEC15-1	CD132	ILEUM	11,7
AIEC15-2			
AIEC16-1	CD132	ILEUM	11,7
AIEC16-2			
LF82			
LF82-Δ*Dsba*			
AIEC08	C142	COLON	2
AIEC09	CD122	COLON	1
AIEC10	C126	ILEUM	1
AIEC11	CD132	ILEUM	5,85
AIEC12	CD89	COLON	1
AIEC17	CD111	COLON / ILEUM	5 / 1
AIEC19	C119	COLON	4,7
AIEC20	CD110	ILEUM	2
AIEC21	CD77	COLON	1
AIEC22	CD61	COLON	1
AIEC25	CD53	COLON	10
AIEC23	CD123	ILEUM	1
AIEC24	CD122	COLON	1
Non *E.coli*	CD53	COLON	–

Figure 21. Consensus UPGMA dendrogram generated from Dice coefficients of *SpeI* and *XbaI* PFGE profiles of the 22 AIEC clonal types detected (AIECtype). * Estimated value calculated after confirmation of putative AIEC detected by the qualitative 96-well plate assay: (Num. of AIEC isolates belonging to AIECtype x / num. of AIEC isolates analyzed by PFGE)*Total AIEC isolates found in a given patient.

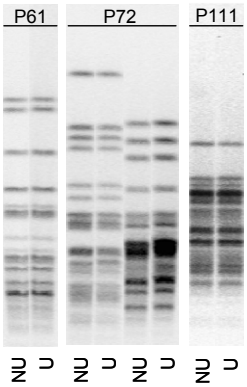
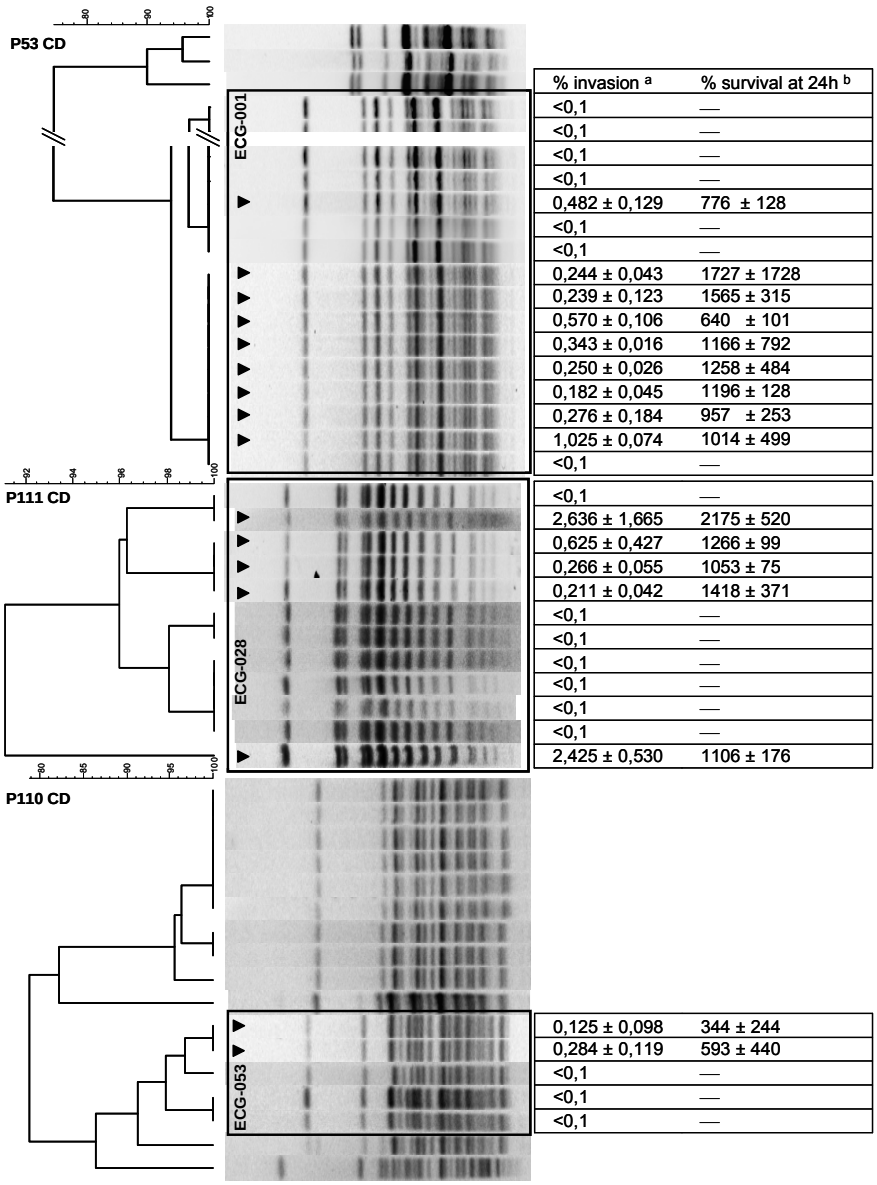


Figure 22. PFGE fingerprints of the *E. coli* subtypes isolated from ulcerated (U) and non ulcerated (NU) mucosa of three CD patients.



% invasion [a]	% survival at 24h [b]
<0,1	—
<0,1	—
<0,1	—
<0,1	—
0,482 ± 0,129	776 ± 128
<0,1	—
<0,1	—
0,244 ± 0,043	1727 ± 1728
0,239 ± 0,123	1565 ± 315
0,570 ± 0,106	640 ± 101
0,343 ± 0,016	1166 ± 792
0,250 ± 0,026	1258 ± 484
0,182 ± 0,045	1196 ± 128
0,276 ± 0,184	957 ± 253
1,025 ± 0,074	1014 ± 499
<0,1	—
<0,1	—
2,636 ± 1,665	2175 ± 520
0,625 ± 0,427	1266 ± 99
0,266 ± 0,055	1053 ± 75
0,211 ± 0,042	1418 ± 371
<0,1	—
<0,1	—
<0,1	—
<0,1	—
<0,1	—
<0,1	—
2,425 ± 0,530	1106 ± 176
0,125 ± 0,098	344 ± 244
0,284 ± 0,119	593 ± 440
<0,1	—
<0,1	—
<0,1	—

Figure 23. Frequently, some *E. coli* isolates with identical pulsotype* were identified as AIEC (indicated by triangles), and others as non-AIEC. *a* Percentage of inoculum surviving after 1 hour of gentamicin treatment; *b* Percentage of intracellular bacteria at 24 hours postinfection relative to the number after 1 hour of gentamicin treatment, defined as 100%.

Discussion

A large number of mucosa-associated *E. coli* isolates per patient were analyzed in the present study to allow for a better comparison of mucosa-adhering bacteria between controls and IBD patients. By determining whether they belonged to the AIEC group, it was possible to compare the pathogenic features of a large collection of AIEC and non-AIEC subtypes. New information about the ecological parameters of the whole *E. coli* population and AIEC pathovar is provided for healthy subjects and CD patients. Virulence genotyping of the twenty-two different AIEC subtypes collected contributes to a better description of the AIEC pathovar.

Table 14. Molecular characterisation of the AIEC collection obtained from 16 Crohn's disease (CD) patients and 6 control subjects (C) including serotype, virulence gene carriage and phylogroup.

AIEC type	Patient	Serotype	Phylo	ExPEC like	*papC*	*sfa/foc DE*	*afa/draBC*	*fimH*	*fimAv*$_{MT78}$	*neuC*	*iucD*	*ibeA*	*hlyA*	*cnf1*	*cdtB*
AIEC04	C 125	O6:HNT	B2	+	+	+	–	+	+	–	+	–	+	+	–
AIEC06	C 128	O6:H5	B2	+	+	+	–	+	–	–	–	–	+	+	+
AIEC08	C 142	O25:H4	B2	+	+	–	–	+	–	–	+	+	–	–	+
AIEC10	C 126	O159:H34	A	–	–	–	–	+	–	–	–	–	–	–	–
AIEC19	C 119	ONT:H–	A	+	–	–	–	+	+	–	+	–	–	–	–
AIEC07	C 124 C 125	O22:H7	B1 D	+	+	–	–	+	–	–	+	–	–	–	–
AIEC01	CD 120	O6:H1	B2	+	–	+	–	+	–	–	+	–	–	–	–
AIEC02	CD 114	O8:H21	B2	+	+	–	–	+	–	+	–	–	–	–	–
AIEC05	CD 114	O1:H–	B2	+	+	–	–	+	–	+	–	–	–	–	–
AIEC09	CD 122	ONT:H–	B2	+	+	+	–	+	+	–	+	–	+	+	–
AIEC11	CD 132	O22:H1	B2	+	–	–	+	+	–	–	+	–	–	–	–
AIEC12	CD 89	O26:H–	B2	+	+	–	+	+	–	–	+	–	–	–	–
AIEC14-1	CD 132	O22:H1	B2	–	–	–	–	+	–	–	–	–	–	–	–
AIEC15-1	CD 132	O22:H1	B2	–	–	–	–	+	–	–	–	–	–	–	–
AIEC16-2	CD 132	O22:H1	B2	–	–	–	–	+	–	–	–	–	–	–	–
AIEC17	CD 111	ONT:HNT	D	+	–	–	–	+	–	+	–	+	–	–	–
AIEC20	CD 110	O11:H18	D	+	+	–	+	+	–	–	+	–	–	–	–
AIEC21	CD 77	O6:H1	B2	+	+	–	–	+	–	–	+	–	+	+	–
AIEC22	CD 61	O119:H21	D	–	–	–	–	+	–	–	+	–	–	–	–
AIEC23	CD 123	O5:HNT	A	–	–	–	–	+	–	–	–	–	–	–	–
AIEC24	CD 122	ONT:H–	A	+	–	–	–	+	+	–	+	–	–	–	–
AIEC25	CD 53	O6:H31	B2	+	+	–	–	+	–	–	–	–	+	+	–

Quantitative-PCR revealed higher *E. coli* counts in I-CD patients in comparison to C-CD, IC-CD and controls, which is in agreement with previous studies [19, 37]; these results were irrespective of the zone sampled along the bowel. However, no difference in *E. coli* diversity was found between CD patients and controls. In fact, *E. coli* subtypes were found to be host-specific, and the same clones were associated with the ulcerated and non-ulcerated mucosa of CD patients. Although a genetic relationship among *E. coli* isolated from CD patients has been already described by ribotyping [38], no genetic relatedness was observed by PFGE in this study, in agreement with the recent results of Sasaki *et al.* [24]. Moreover, phylogenetic groups A, B1, B2, and D were equally

distributed in CD patients and healthy subjects, with B2 being the most abundant. The majority of mucosa-associated *E. coli* from both CD patients and controls showed "uropathogenic" features, which are characteristic of B2 and D phylogroups [39-41]. These features have already been described for the resident colonic microbiota in normal mucosa [42] and are thought to possess a fitness or colonisation function [43].

Table 15. Virulence-associated genes frequencies and phylogroup distribution of AIEC and non-AIEC strains isolated from Crohn's disease (CD) patients and controls (C).

Virulence gene/ Phylogroup	AIEC (n=22)	non-AIEC (n=38)	P value	CD- AIEC (n=16)	C- AIEC (n=6)	P value
ExPEC-like	73	58	NS	83	69	NS
papC	50	47	NS	44	67	NS
sfa/focDE	18	18	NS	13	33	NS
afa/draBC	14	11	NS	19	0	NS
fimH	100	92	NS	100	100	NS
$fimAv_{MT78}$	18	21	NS	13	33	NS
neuC-K1	14	16	NS	19	0	NS
iucD	55	58	NS	50	67	NS
ibeA	9	11	NS	6	17	NS
hlyA	23	18	NS	19	33	NS
cnf1	23	18	NS	19	33	NS
cdtB	9	5	NS	0	33	NS
DEC (EAEC)	0	3	NS	0	0	NS
A	18	13	NS	13	33	NS
B1	5	18	NS	0	17	NS
B2	64	38	0.044	69	50	NS
D	14	30	NS	19	0	NS
ATYPICAL	0	3	NS	0	0	-

Values expressed in percentages. ExPEC-like: strains with ≥2 virulence genes typical from extraintestinal pathogenic *E. coli*. EAEC: Enteroaggregative *E. coli*.

The distribution of virulence genes between *E. coli* subtypes from CD patients and controls was very similar, except for *iucD*, which was found less frequently in CD-*EC* subtypes. This gene is involved in the biosynthesis of the siderophore aerobactin. Since active CD patients present impaired intestinal iron absorption at the intestinal level [44], iron availability will likely increase in the intestinal mucosa. Therefore, *E. coli* in CD could lose portions of their iron-absorption-related genetic pool without any ecologic cost. This change in iron availability could also potentially drive changes in the entire composition of the microbial population.

The AIEC prevalence was higher in CD patients than in controls in both the ileum and colon. Moreover, an increase in AIEC prevalence was observed in comparison with previous studies [19, 22, 23], probably due to our methodological approach, which enabled us to obtain a more accurate prevalence value by analyzing a greater number of *E. coli* per patient.

In the present study, we assessed the abundance of AIEC within the *E. coli* population, an analysis that has never been performed in any studies reported thus far. Although the mean AIEC abundance was low and variable, AIEC was more abundant in CD patients than in controls, and this difference was greater for Crohn's ileitis patients. This could be partially explained by an increase in CEACAM6 receptors in the ileal mucosa of CD patients, as described by Barnich *et al.* [45]. However, while we observed a higher prevalence and abundance in both the ileum and colon, Barnich *et al.* observed that colonic specimens expressed CEACAM6 at similar levels in CD patients and controls. The factors determining AIEC colonisation in the colonic mucosa of CD patients could be diverse. Upstream colonisation of the ileum might be a cause, in the sense that the composition of the ileal microbial population would determine the colonic mucosa. However, other host and/or environmental factors could be involved.

The greater richness of AIEC in CD patients suggests that this pathovar might be more permanent in CD intestinal mucosa owing to better host-environmental conditions. Greater persistence in the environment could eventually contribute to the diversification of clones. In particular, patient CD132 harboured up to 8 distinct AIEC clones, 7 of which were grouped within the same PFGE cluster (Figure 21), indicating minor changes that could be attributed to clonal turnover [46].

The frequency and nature of genetic virulence determinants were similar between AIEC and non-AIEC subtypes. Moreover, both *E. coli* types exhibited extraintestinal pathogenic features. Phylogroup B2 was more prevalent in AIEC subtypes than in non-AIEC. As expected, *fimH* was present in all AIEC subtypes, and *papC* was widespread within the pathovar, which is in accordance with previous studies [19, 23] but in contrast to others [24]. No CD-associated *E. coli* was positive for *cnf1*/*hlyA* in previous studies [47], whereas 23% of the AIEC subtypes from our CD population were positive for this gene. In addition, known invasive determinants of APEC (*ibeA*) already described for AIEC [19] were present in 9% of our AIEC collection, as well as in the LF82 AIEC reference strain. As revealed by other investigators [19, 23, 38, 47, 48], these results suggest that AIEC could be a novel group of ExPEC associated with IBDs, distinguishable from the mucosa-associated *E. coli* of patients with colorectal cancer and healthy subjects. It would be interesting in further work to identify AIEC properties within a collection of the most important ExPEC strains and compare their virulence genotypes more extensively. The definition of virulence within ExPEC remains unclear. Recently, genes associated with the virulence of B2 strains have been speculated to be involved in complex host-commensal niche

colonisation, with virulence arising as a by-product [49, 50]. The presence of AIEC in healthy subjects suggests that AIEC are facultative pathogens that can cause disease in susceptible hosts.

Interestingly, several *E. coli* isolates belonging to the same clonal group showed different invasive properties (Figure 23). This could be explained in several ways, for example, differences in gene expression, the existence of single nucleotide polymorphisms not detectable by PFGE, or the loss or gain of DNA by horizontal gene transfer.

In conclusion, greater AIEC prevalence, abundance, and richness are present in CD patients, supporting the hypothesis that this pathovar might play an etiologic role in Crohn's disease. Molecular characterisation of this pathovar revealed great internal variability of serotypes, genotypes, and pulsotypes* sharing pathogenic features of extraintestinal pathogenic *E. coli*. In contrast, phylogeny was less variable, with the B2 phylogroup exhibiting greater prevalence. Further research on the differences between clonally identical isolates and different invasive properties could provide useful information about the genes characteristic of AIEC.

References

1 Baumgart DC, Carding SR. Inflammatory bowel disease: cause and immunobiology. The Lancet. 2007;369:1627-40.

2 Halfvarson J, Bodin L, Tysk C, *et al.* Inflammatory bowel disease in a Swedish twin cohort: a long-term follow-up of concordance and clinical characteristics. Gastroenterology. 2003;124:1767-73.

3 Sartor R. Mechanisms of Disease: pathogenesis of Crohn's disease and ulcerative colitis. Nat Clin Pract Gastroenterol Hepatol. 2006;3:390-407.

4 Strober W, Fuss I, Mannon P. The fundamental basis of inflammatory bowel disease. J Clin Invest. 2008;117:514-21.

5 Xavier RJ, Podolsky DK. Unravelling the pathogenesis of inflammatory bowel disease. Nature. 2007;448:427-34.

6 De Hertogh G, Aerssens J, Geboes KP, *et al.* Evidence for the involvement of infectious agents in the pathogenesis of Crohn's disease. World J Gastroenterol. 2008 14:845-52.

7 Rutgeerts P, Hiele M, Geboes K, *et al.* Controlled trial of metronidazole treatment for prevention of Crohn's recurrence after ileal resection. Gastroenterology. 1995;108:1617-21.

8 Rutgeerts P, Geboes K, Peeters M, *et al.* Effect of faecal stream diversion on recurrence of Crohn's disease in the neoterminal ileum. Lancet. 1991;338:771-4.

9 Sartor RB. Microbial Influences in Inflammatory Bowel Diseases. Gastroenterology. 2008;134:577-94.

10 Sellon RK, Tonkonogy S, Schultz M*, et al.* Resident Enteric Bacteria Are Necessary for Development of Spontaneous Colitis and Immune System Activation in Interleukin-10-Deficient Mice. Infect Immun. 1998;66:5224-31.

11 Stephen B. Hanauer M. Inflammatory Bowel Disease: Epidemiology, Pathogenesis, and Therapeutic Opportunities. Inflamm Bowel Dis. 2006;12:S3-S9.

12 Mangin I, Bonnet R, Seksik P*, et al.* Molecular inventory of faecal microflora in patients with Crohn's disease. FEMS Microbiol Ecol. 2004;50:25-36.

13 Seksik P, Rigottier-Gois L, Gramet G*, et al.* Alterations of the dominant faecal bacterial groups in patients with Crohn's disease of the colon. Gut. 2003;52:237-42.

14 Martinez-Medina M, Aldeguer X, Gonzalez-Huix F *et al.* Abnormal microbiota composition in the ileocolonic mucosa of Crohn's disease patients as revealed by polymerase chain reaction-denaturing gradient gel electrophoresis. Inflamm Bowel Dis. 2006;12:1136-45.

15 Ott SJ, Musfeldt M, Wenderoth DF*, et al.* Reduction in diversity of the colonic mucosa associated bacterial microflora in patients with active inflammatory bowel disease. Gut. 2004;53:685-93.

16 Sokol H, Lepage P, Seksik P*, et al.* Molecular comparison of dominant microbiota associated with injured versus healthy mucosa in ulcerative colitis. Gut. 2007;56:152-a-4.

17 Bibiloni R, Mangold M, Madsen KL*, et al.* The bacteriology of biopsies differs between newly diagnosed, untreated, Crohn's disease and ulcerative colitis patients. J Med Microbiol. 2006;55:1141-9.

18 Gophna U, Sommerfeld K, Gophna S*, et al.* Differences between Tissue-Associated Intestinal Microfloras of Patients with Crohn's Disease and Ulcerative Colitis. J Clin Microbiol. 2006;44:4136-41.

19 Baumgart M, Dogan B, Rishniw M*, et al.* Culture independent analysis of ileal mucosa reveals a selective increase in invasive *Escherichia coli* of novel phylogeny relative to depletion of *Clostridiales* in Crohn's disease involving the ileum. ISME J. 2007;1:403-18.

20 Salzman NH, Underwood MA, Bevins CL. Paneth cells, defensins, and the commensal microbiota: A hypothesis on intimate interplay at the intestinal mucosa. Semin Immunol. 2007;19:70-83.

21 Lupp C, Robertson ML, Wickham ME*, et al.* Host-Mediated Inflammation Disrupts the Intestinal Microbiota and Promotes the Overgrowth of *Enterobacteriaceae*. Cell Host Microbe. 2007;2:119-29.

22 Darfeuille-Michaud A, Boudeau J, Bulois P*, et al.* High prevalence of adherent-invasive *Escherichia coli* associated with ileal mucosa in Crohn's disease. Gastroenterology. 2004;127:412-21.

23 Martin HM, Campbell BJ, Hart CA*, et al.* Enhanced *Escherichia coli* adherence and invasion in Crohn's disease and colon cancer. Gastroenterology. 2004;127:80-93.

24 Sasaki M, Sitaraman SV, Babbin BA*, et al.* Invasive *Escherichia coli* are a feature of Crohn's disease. Lab Invest. 2007;87:1042-54.

25 Barnich N, Darfeuille-Michaud A. Adherent-invasive *Escherichia coli* and Crohn's disease. Curr Opin Gastroenterol. 2007;23:16-20.

26 Eaves-Pyles T, Allen CA, Taormina J*, et al. Escherichia coli* isolated from a Crohn's disease patient adheres, invades, and induces inflammatory responses in polarized intestinal epithelial cells. Int J Med Microbiol. 2008;298:397-409.

27 Rolhion N, Darfeuille-Michaud A. Adherent-invasive *Escherichia coli* in inflammatory bowel disease. Inflamm Bowel Dis. 2007;13:1277-83.

28 Huijsdens XW, Linskens RK, Mak M, *et al.* Quantification of Bacteria Adherent to Gastrointestinal Mucosa by Real-Time PCR. J Clin Microbiol. 2002;40:4423-7.

29 Thompson CJ, Daly C, Barrett TJ, *et al.* Insertion Element IS3-Based PCR Method for Subtyping *Escherichia coli* O157:H7. J Clin Microbiol. 1998;36:1180-4.

30 CDC PulseNet [database online]. One-Day (24-28 h) Standardized Laboratory Protocol for Molecular Subtyping of *Escherichia coli* O157:H7, non-typhoidal *Salmonella* serotypes, and *Shigella sonnei* by Pulsed Field Gel Electrophoresis (PFGE). Centers for Disease Control and Prevention; Accessed September 2004.

31 Clermont O, Bonacorsi S, Bingen E. Rapid and Simple Determination of the *Escherichia coli* Phylogenetic Group. Appl Environ Microbiol. 2000;66:4555-8.

32 Blanco M, Blanco J, Dahbi G, *et al*. Identification of two new intimin types in atypical enteropathogenic *Escherichia coli*. Int Microbiol. 2006;9:103-10.

33 Naves P, del Prado G, Huelves L, *et al.* Correlation between virulence factors and in vitro biofilm formation by *Escherichia coli* strains. Microb Pathog. 2008;45:86-91.

34 Guinée PA, Agterberg CM, Jansen WH. *Escherichia coli* O antigen typing by means of a mechanized microtechnique. Appl Microbiol. 1972;24:127-31.

35 Boudeau J, Glasser A-L, Masseret E, *et al.* Invasive Ability of an *Escherichia coli* Strain Isolated from the Ileal Mucosa of a Patient with Crohn's Disease. Infect Immun. 1999;67:4499-509.

36 Glasser A-L, Boudeau J, Barnich N, *et al.* Adherent Invasive *Escherichia coli* Strains from Patients with Crohn's Disease Survive and Replicate within Macrophages without Inducing Host Cell Death. Infect Immun. 2001;69:5529-37.

37 Fujita H, Eishi Y, Ishige I, *et al.* Quantitative analysis of bacterial DNA from *Mycobacteria* spp., *Bacteroides vulgatus*, and *Escherichia coli* in tissue samples from patients with inflammatory bowel diseases. J Gastroenterol. 2002;37:509-16.

38 Masseret E, Boudeau J, Colombel JF, *et al.* Genetically related *Escherichia coli* strains associated with Crohn's disease. Gut. 2001;48:320-5.

39 Boyd EF, Hartl DL. Chromosomal Regions Specific to Pathogenic Isolates of *Escherichia coli* Have a Phylogenetically Clustered Distribution. J Bacteriol. 1998;180:1159-65.

40 Johnson JR, Oswald E, O'Bryan TT, *et al.* Phylogenetic Distribution of Virulence-Associated Genes among *Escherichia coli* Isolates Associated with Neonatal Bacterial Meningitis in The Netherlands. J Infect Dis. 2002;185:774-84.

41 Picard B, Garcia JS, Gouriou S, *et al.* The Link between Phylogeny and Virulence in *Escherichia coli* Extraintestinal Infection. Infect Immun. 1999;67:546-53.

42 Nowrouzian FL, Adlerberth I, Wold AE. Enhanced persistence in the colonic microbiota of *Escherichia coli* strains belonging to phylogenetic group B2: role of virulence factors and adherence to colonic cells. Microbes Infect. 2006;8:834-40.

43 Nowrouzian F, Adlerberth I, Wold AE. P fimbriae, capsule and aerobactin characterize colonic resident Escherichia coli. Epidemiol Infect 2001;126:11-8.

44 Semrin G, Fishman DS, Bousvaros A, *et al*. Impaired intestinal iron absorption in Crohn's disease correlates with disease activity and markers of inflammation. Inflamm Bowel Dis. 2006;12:1101-6.

45 Barnich N, Carvalho F, Glasser AL, *et al*. CEACAM6 acts as a receptor for adherent-invasive *E. coli*, supporting ileal mucosa colonization in Crohn disease. J Clin Invest. 2007;117:1566-74.

46 Karch H, Russmann H, Schmidt H, *et al.* Long-term shedding and clonal turnover of enterohemorrhagic *Escherichia coli* O157 in diarrheal diseases. J Clin Microbiol. 1995;33:1602-5.

47 Bronowski C, Smith SL, Yokota K, *et al.* A subset of mucosa-associated *Escherichia coli* isolates from patients with colon cancer, but not Crohn's disease, share pathogenicity islands with urinary pathogenic *E. coli*. Microbiology. 2008;154:571-83.

48 Kotlowski R, Bernstein CN, Sepehri S, *et al.* High prevalence of *Escherichia coli* belonging to the B2+D phylogenetic group in inflammatory bowel disease. Gut. 2007;56:669-75.

49 Le Gall T, Clermont O, Gouriou S, *et al.* Extraintestinal Virulence Is a Coincidental By-Product of Commensalism in B2 Phylogenetic Group *Escherichia coli* Strains. Mol Biol Evol. 2007;24:2373-84.

50 Levin B. The evolution and maintenance of virulence in microparasites. Emerg Infect Dis. 1996;2:93-102.

Chapter 3

Characterisation of adherent-invasive *E. coli* strains isolated from the intestinal mucosa of Crohn's disease patients and control subjects

3.1

Biofilm formation as a novel phenotypic feature of adherent- invasive *Escherichia coli* (AIEC)

Introduction

Crohn's disease (CD) is a chronic-relapsing inflammatory bowel disease (IBD) that can affect the entire gastrointestinal tract. The incidence rate varies from 1 to 20 cases per 10^5 people per year and is still rising in some countries [1]. Although the aetiology of CD remains elusive to date, it is widely accepted that several factors are involved in the onset or perpetuation of the disease. These factors include genetic and immunologic features that confer host susceptibility, and external or environmental factors such as microorganisms and lifestyle [2, 3]. Environmental factors play an important role because there is a low concordance between identical twins, both for CD and ulcerative colitis (UC) [4]. The involvement of microbes in the onset or perpetuation of inflammation has been extensively studied [5-10]. To date, some pathogens have been proposed as causative agents. In particular, adherent-invasive *E. coli* (AIEC) is increasing in relevance because it has been reported to be more prevalent in CD patients than in controls in several countries (France [11], United Kingdom [12], USA [13, 14], and Spain [15]). AIEC strains have the ability to adhere to and to invade intestinal epithelial cells *in vitro* as well as to survive and replicate within macrophages without inducing host-cell death and promoting tumour necrosis factor (TNF) α release. No unique genetic sequences have been described for AIEC, nor have specific genes of diarrhoeagenic pathovars been detected yet for AIEC, but they do carry many virulence-associated genes characteristic of extraintestinal pathogenic *E. coli* (ExPEC) [13, 15, 16]. For that reason, AIEC pathovar has been speculated to be closely related to ExPEC pathovar.

In a previous work, we observed that some CD patients showed a high diversity of AIEC subtypes associated to their intestinal mucosa [15]. In a given patient, we could detect up to 8 different clones as assessed by pulsed field gel electrophoresis. On the other hand, AIEC abundance, richness and diversity were lower in non-IBD controls. We hypothesized that the higher diversification of clones could be explained by a long-term colonization of AIEC in CD. Biofilm formation can be a way to persistently colonize the intestinal mucosa [17], as has been reported for commensal microbiota in healthy subjects [18]. Moreover, for

certain organisms such as the species belonging to the uropathogenic *E. coli* pathovar (UPEC) – which form intracellular biofilms [19] – and to the enteroaggregative pathogenic *E. coli* pathovar (EAEC) – which form thick biofilms that adhere to the apical side of enterocytes [20]–, active biofilm formation is feature of their pathogenesis. For that reason, the primarily aim of this work was to determine the biofilm formation capacity of AIEC strains and non-AIEC strains, both isolated from the intestinal mucosa.

We herein report a new phenotypic feature of the recently described AIEC pathovar which is the ability to form biofilms *in vitro*. In addition, we illustrate those seropathotypes and phylotypes more frequently found amongst biofilm producers.

Materials and methods

Bacterial strains

Amongst the collection of 65 *E. coli* strains, sixty-one (93.8%) were isolated from human intestinal mucosa in previous studies [15, 21]. In particular, 35 strains (16 of them were AIEC) came from CD patients, one (which belonged to AIEC pathovar) came from a patient suffering from ulcerative colitis, and 25 (of which 6 were AIEC) came from non-IBD controls. Also included were four additional AIEC strains that came from patients with extraintestinal infection (two with sepsis and two with urinary tract infection [22, 23]). AIEC reference strain LF82 and the isogenic mutant LF82-Δ*fliC* were used as controls. Relevant characteristics of the strains that were known prior to this study are compiled in Table 16.

All procedures were approved by the ethics committee of clinical investigation of the Hospital Josep Trueta of Girona in compliance with the Helsinki declaration.

Biofilm formation assay

Biofilm formation assays were performed using a previously described method [24] with some modifications [25]. Strains were grown overnight in Luria-Bertani broth with 5 g l^{-1} of glucose (Sigma-Aldrich, St. Louis, USA) at 35.5ºC, then 1/100 dilutions were made in M63 minimal medium (US Biological, Swampscott, USA) supplemented with 8 g l^{-1} (0.8%) glucose. Then, 130-μl aliquots were placed in wells of non-cell-treated polystyrene microtiter plates (Greiner Bio-one, Stuttgart, Germany) and incubated overnight at 30ºC without shaking. Afterwards, growth optical densities (OD) were read at 630 nm; then the wells

were washed once, adhered bacteria were stained with 1% crystal violet solubilised in ethanol, and ODs read at 570 nm. Biofilm measurements were calculated using the formula *SBF = (AB-CW)/G*, in which *SBF* is the specific biofilm formation, *AB* is the OD_{570nm} of the attached and stained bacteria, *CW* is the OD_{570nm} of the stained control wells containing only bacteria-free medium (to eliminate unspecific or abiotic OD values), and *G* is the OD_{630nm} of cell growth in broth [26, 27]. For each assay, 16 wells per strain were analyzed, and the assays were performed in triplicate, which resulted in a total of 48 wells per each tested strain and control. The degree of biofilm production was classified in three categories: weak (SBF≤0.5), moderate (0.5>SBF≤1), and strong (SBF>1).

Adhesion and invasion assays in epithelial cells Intestine-407

Adhesion (I_ADH) and invasion indices (I_INV) were determined as explained in chapter 2 (quantitative analysis) following the gentamicin protection assay previously described elsewhere [21].

Survival and replication in macrophages J774

Intra-macrophage replication index (I_REPL) was performed as explained in chapter 2, following the method previously described elsewhere [11, 28]. Those strains with I_INV > 0.1 and I_REPL > 100% were classified as AIEC in this study (AIEC-phenotype).

Serotyping

Determination of O and H antigens was carried out using the method previously described by Guinée *et al.* [29]. Strains which failed to achieve motility on semisolid medium were considered nonmotile and designated H–.

Phylotyping and virulence genotyping by PCR

Determination of the major *E. coli* phylogenetic group (A, B1, B2, and D) was performed as previously described by Clermont *et al.* [30]. Virulence genotyping was performed as explained in chapter 2 following the methods described elsewhere [25, 31].

Table 16. Phenotypic and genotypic characteristics of the bacterial strains used in this study.

Id Strain	Origin	AIEC phenotype	Serotype	Phylogroup	Virulence gene carriage
AIEC19	C	+	ONT: H–	A	*iucD, fimH, fimAv$_{MT78}$*
AIEC07	C	+	O22: H7	B1	*papC, iucD, fimH*
AIEC04	C	+	O6: HNT	B2	*papC, sfa/focDE, iucD, hlyA, cnf1, fimH, fimAv$_{MT78}$*
AIEC10	C	+	O159: H34	A	*fimH*
AIEC06	C	+	O6: H5	B2	*papC, sfa/focDE, hlyA, cnf1, fimH, cdtB*
AIEC08	C	+	O25:H4	B2	*papC, iucD, ibeA, fimH, cdtB*
AIEC25	CD	+	O6: H31	B2	*papC, hlyA, cnf1, fimH*
AIEC21	CD	+	O6: H1	B2	*papC, iucD, hlyA, cnf1, fimH*
AIEC12	CD	+	O26: H–	B2	*papC, afa/draBC, iucD, fimH*
AIEC20	CD	+	O11: H18	D	*papC, afa/draBC, iucD, fimH*
AIEC17	CD	+	ONT: HNT	D	*neuC, ibeA, fimH*
AIEC05	CD	+	O1: H–	B2	*papC, neuC, fimH*
AIEC02	CD	+	O8: H21	B2	*papC, neuC, fimH*
AIEC01	CD	+	O6: H1	B2	*sfa/focDE, iucD, fimH*
AIEC09	CD	+	ONT: H–	B2	*papC, sfa/focDE, iucD, hlyA, cnf1, fimH, fimAv$_{MT78}$*
AIEC24	CD	+	ONT: H–	A	*iucD, fimH, fimAv$_{MT78}$*
AIEC23	CD	+	O5: HNT	A	*fimH*
AIEC11	CD	+	O22: H1	B2	*afa/draBC, iucD, fimH*
AIEC15-1	CD	+	O22: H1	B2	*fimH*
AIEC14-1	CD	+	O22: H1	B2	*fimH*
AIEC16-2	CD	+	O22: H1	B2	*fimH*
LF82	CD	+	O83: H1	B2	*ibeA, fimH, fimAv$_{MT78}$*
AIEC13	UC	+	O25: H4	B2	*papC, iucD, hlyA, cnf1, ibeA, fimH, cdtB*
PP16	SEPSIS	+	O83: H1	B2	*ibeA, fimH, fimAv$_{MT78}$*
FV7563	UTI	+	O25: H4	B2	*afa/draBC, iucD, fimH*
OL96A	UTI	+	O6: H1	B2	*papC, sfa/focDE, iucD, hlyA, cnf1, fimH*
PP215	SEPSIS	+	O6: H1	B2	*papC, sfa/focDE, iucD, hlyA, cnf1, fimH*
ECG-046	C	–	ONT: H29	B1	*iucD*, pCDV432
ECG-060	C	–	O102: H6	D	*papC, iucD, fimH*
ECG-037	C	–	O1: H–	D	*papC, neuC, iucD, fimH, fimAv$_{MT78}$*
ECG-016	C	–	O55: H–	A	*neuC, iucD, fimH*
ECG-017	C	–	O22: H18	B2	*sfa/focDE, fimH*
ECG-022	C	–	O15: H–	A	*afa/draBC, neuC, iucD, fimH*
ECG-043	C	–	O83: H1	B2	*ibeA, fimH, fimAv$_{MT78}$*
ECG-041	C	–	O2: H–	B2	*papC, sfa/focDE, iucD, hlyA, cnf1, fimH, cdtB*
ECG-012	C	–	O18: H1	B2	*papC, sfa/focDE, iucD, hlyA, cnf1, fimH*
ECG-025	C	–	O6: H31	B2	*papC, hlyA, cnf1, fimH*
ECG-049	C	–	O15: H16	B2	*papC, iucD, fimH*
ECG-031	C	–	O161: H4	D	*iucD, fimH*
ECG-023	C	–	ONT: H–	Atypical	*iucD, fimH, fimAv$_{MT78}$*
ECG-054	C	–	O14: H–	D	*papC, iucD, fimH*
ECG-008	C	–	O14: H7	B2	*papC, iucD, ibeA, fimH, fimAv$_{MT78}$*
ECG-004	C	–	O22: H7	B1	*fimH*
ECG-013	C	–	O18: H1	B2	*papC, sfa/focDE, iucD, hlyA, cnf1, fimH*
ECG-055	C	–	O17: H18	D	*papC, iucD, fimH*
ECG-024	C	–	O174: H–	B1	*fimH*
ECG-064	CD	–	O166: H21	B1	*fimH*
ECG-042	CD	–	O7: H6	B2	*fimH, fimAv$_{MT78}$*
ECG-001	CD	–	O6: H31	B2	*papC, hlyA, cnf1, fimH*
ECG-005	CD	–	O4: H5	B2	*papC, sfa/focDE, hlyA, cnf1, fimH*
ECG-065	CD	–	O141: H–	A	*fimH, fimAv$_{MT78}$*
ECG-047	CD	–	O119: H21	D	*iucD, fimH*
ECG-019	CD	–	ONT: H–	A	
ECG-018	CD	–	ONT: H–	A	*iucD*
ECG-002	CD	–	O175: H–	B1	*fimH*
ECG-034	CD	–	O127: H28	D	*iucD, fimH*
ECG-021	CD	–	O54: H21	B1	*fimH*
ECG-063	CD	–	ONT: H21	B1	*fimH*
ECG-056	CD	–	ONT: H18	D	*papC, afa/draBC, iucD, fimH*
ECG-057	CD	–	O11: H18	D	*papC, afa/draBC, iucD, fimH*
ECG-053	CD	–	O11: H–	D	*papC, afa/draBC, iucD, fimH*
ECG-059	CD	–	ONT: H18	D	*papC, iucD, fimH*
ECG-026	CD	–	O1: H–	B2	*papC, neuC,ibeA, fimH, fimAv$_{MT78}$*
ECG-015	CD	–	O2: H6	B2	*papC, sfa/focDE, neuC, hlyA, cnf, fimH*
ECG-009	CD	–	O83: H1	B2	*sfa/focDE, neuC, iucD, ibeA, fimH, fimAv$_{MT78}$, cdtB*

Abbreviators: CD: Crohn's disease; UC: ulcerative colitis; C: non-IBD control, UTI: urinary tract infection. AIEC phenotype +: strains that adhere to and invade Intestine-407 cells and that were able to survive and/or replicate within J774 macrophages in vitro. ONT: serogroup non-typeable, HNT: flagellar antigen non-typeable. H– : nonmotile strain.

Table 17. AIEC phenotype and category of biofilm formation ability of the strains.

Id Strain	AIEC phenotype	I_ADH (mean ± SD)		I_INV (mean ± SD)		I_REPL (mean ± SD)		SBF (mean ± SD)		Biofilm formation category
AIEC19	+	2.40	± 0.65	0.111	± 0.016	1568.1	± 1726	0.053	± 0.019	W
AIEC07	+	20.00	± 13.50	0.565	± 0.392	1692.6	± 296.8	2.391	± 0.317	S
AIEC04	+	21.50	± 9.00	0.320	± 0.016	584.7	± 418.5	0.772	± 0.132	M
AIEC10	+	6.00	± 0.98	0.226	± 0.192	1413.7	± 51.4	0.610	± 0.175	M
AIEC06	+	10.25	± 3.25	0.177	± 0.019	1717.7	± 307.9	1.212	± 0.233	S
AIEC08	+	1.13	± 0.18	0.172	± 0.066	104.8	± 49.7	0.397	± 0.133	W
AIEC25	+	2.75	± 1.33	0.482	± 0.129	775.9	± 128.3	0.437	± 0.129	W
AIEC21	+	17.00	± 7.75	0.109	± 0.013	1297.1	± 625.2	0.558	± 0.205	M
AIEC12	+	22.25	± 4.00	0.142	± 0.017	193.7	± 55.9	0.125	± 0.052	W
AIEC20	+	14.25	± 6.25	0.125	± 0.098	343.9	± 244.6	0.284	± 0.116	W
AIEC17	+	21.75	± 17.50	0.266	± 0.055	1053.0	± 75.0	0.840	± 0.286	M
AIEC05	+	9.50	± 2.25	0.202	± 0.042	704.9	± 714.0	0.181	± 0.072	W
AIEC02	+	0.85	± 1.03	0.802	± 0.035	2187.8	± 4.8	0.106	± 0.035	W
AIEC01	+	16.00	± 9.25	0.284	± 0.106	1566.7	± 1060	0.700	± 0.177	M
AIEC09	+	5.25	± 4.00	0.216	± 0.010	2562.3	± 240.6	0.068	± 0.035	W
AIEC24	+	1.98	± 1.40	0.309	± 0.138	1625.6	± 115.6	0.076	± 0.044	W
AIEC23	+	9.75	± 0.70	0.568	± 0.148	2362.1	± 250.2	0.300	± 0.093	W
AIEC11	+	0.83	± 0.19	2.125	± 1.164	739.4	± 477.4	0.537	± 0.129	M
AIEC15-1	+	25.00	± 15.75	2.261	± 1.349	776.9	± 304.8	1.090	± 0.407	S
AIEC14-1	+	4.25	± 3.50	0.508	± 0.081	847.9	± 512.8	0.654	± 0.153	M
AIEC16-2	+	10.00	± 1.425	0.305	± 0.159	659.7	± 437.0	0.502	± 0.134	M
LF82	+	25.00	± 5.25	2.261	± 0.011	776.9	± 252.4	1.641	± 0.326	S
AIEC13	+	1.20	± 4.25	0.104	± 0.000	1045.9	± 181.6	0.772	± 0.211	M
PP16	+	8.00	± 0.98	1.400	± 0.081	225.9	± 541.2	1.012	± 0.268	S
FV7563	+	6.75	± 6.00	0.129	± 0.072	470.0	± 264.0	0.518	± 0.226	M
OL96A	+	5.25	± 5.00	0.388	± 0.159	457.5	± 259.3	1.208	± 0.202	S
PP215	+	0.83	± 0.60	0.453	± 0.350	1425.4	± 229.4	0.546	± 0.139	M
ECG-046	−	-		<0.1		-		0.004	± 0.010	W
ECG-060	−	-		<0.1		-		0.127	± 0.041	W
ECG-037	−	-		<0.1		-		0.042	± 0.024	W
ECG-016	−	-		<0.1		-		0.134	± 0.085	W
ECG-017	−	-		<0.1		-		1.074	± 0.286	S
ECG-022	−	-		<0.1		-		0.143	± 0.090	W
ECG-043	−	-		<0.1		-		1.187	± 0.511	S
ECG-041	−	-		<0.1		-		0.301	± 0.123	W
ECG-012	−	-		<0.1		-		0.741	± 0.259	M
ECG-025	−	-		<0.1		-		0.154	± 0.043	W
ECG-049	−	-		<0.1		-		0.384	± 0.160	W
ECG-031	−	-		<0.1		-		0.067	± 0.024	W
ECG-023	−	0.90	± 0.65	0.052	± 0.003	-		0.038	± 0.020	W
ECG-054	−	-		<0.1		-		0.209	± 0.128	W
ECG-008	−	-		<0.1		-		0.817	± 0.288	M
ECG-004	−	-		<0.1		-		1.113	± 0.234	S
ECG-013	−	-		<0.1		-		0.516	± 0.332	M
ECG-055	−	-		<0.1		-		0.108	± 0.033	W
ECG-024	−	-		<0.1		-		0.037	± 0.016	W
ECG-064	−	-		<0.1		-		0.553	± 0.171	M
ECG-042	−	-		<0.1		-		0.348	± 0.147	W
ECG-001	−	-		<0.1		-		0.299	± 0.106	W
ECG-005	−	-		<0.1		-		0.404	± 0.103	W
ECG-065	−	-		0.061	± 0.070	-		0.026	± 0.022	W
ECG-047	−	1.93	± 1.95	0.259	± 0.084	-		0.007	± 0.016	W
ECG-019	−	-		<0.1		-		0.439	± 0.057	W
ECG-018	−	-		<0.1		-		0.058	± 0.042	W
ECG-002	−	-		<0.1		-		0.039	± 0.023	W
ECG-034	−	-		<0.1		-		0.293	± 0.101	W
ECG-021	−	6.00	± 4.00	0.033	± 0.011	-		0.311	± 0.117	W
ECG-063	−	-		<0.1		-		0.195	± 0.064	W
ECG-056	−	-		<0.1		-		0.124	± 0.047	W
ECG-057	−	11.75	± 7.25	0.013	± 0.011	-		0.241	± 0.094	W
ECG-053	−	-		<0.1		-		0.262	± 0.083	W
ECG-059	−	-		<0.1		-		0.200	± 0.137	W
ECG-026	−	-		<0.1		-		0.418	± 0.189	W
ECG-015	−	5.25	± 2.75	0.038	± 0.004	-		1.035	± 0.219	S
ECG-009	−	-		<0.1		-		1.346	± 0.205	S

Adhesion, invasion, intra-macrophage replication, and biofilm formation indices are specified. Abbreviators: AIEC phenotype +: strains that adhere to and invade Intestine-407 cells and that were able to survive and/or replicate within J774 macrophages in vitro. I_ADH: adhesion index; I_INV: invasion index; I_REPL: replication index; SBF: specific biofilm formation index; W: weak biofilm producer; M: moderate biofilm producer; and S: strong biofilm producer.

Statistical analyses

Quantitative parameters, such as SBF, adhesion, and invasion indices were compared by one-way ANOVA. In cases for which the interaction between several factors was of interest, a factorial ANOVA was applied. Correlation between quantitative variables was assessed by Pearson correlation coefficient. Fisher's exact test (small contingency tables) or Pearson's X^2 tests (frequencies higher than five within cells) were used to measure the significance of frequency values.

Results

AIEC strains are stronger biofilm producers than non-AIEC strains isolated from the intestinal mucosa

The present study involved a collection of 65 *E. coli* strains, 27 of which (41.5%) were classified as AIEC by their ability to adhere to and invade intestinal epithelial cells, and to survive and replicate within macrophages, as previously described [11] (Table 16).

Within the category of weak biofilm producers, 74.4% of strains were non-AIEC, whereas 65.4% of moderate to strong biofilm producers were AIEC (P=0.002). Amongst these AIEC strains, 22.2% were strong biofilm producers, and 40.7% were moderate biofilm producers (Table 17). Similar results were obtained when SBF index values were compared. As shown in Figure 24, the mean SBF index was higher in AIEC strains than in non-AIEC (SBF_{AIEC}=0.65±0.53; $SBF_{NON-AIEC}$=0.36±0.36; P=0.012).

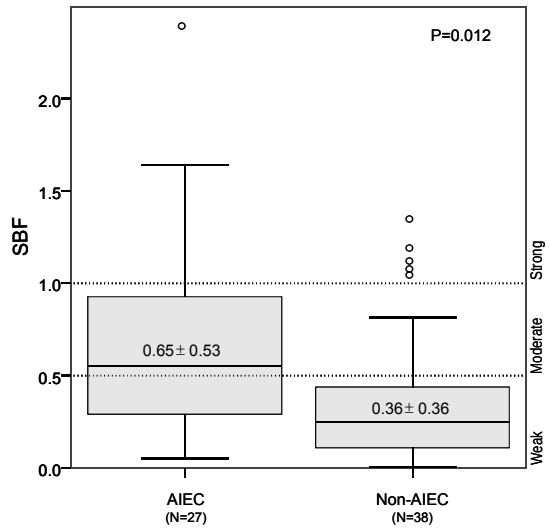
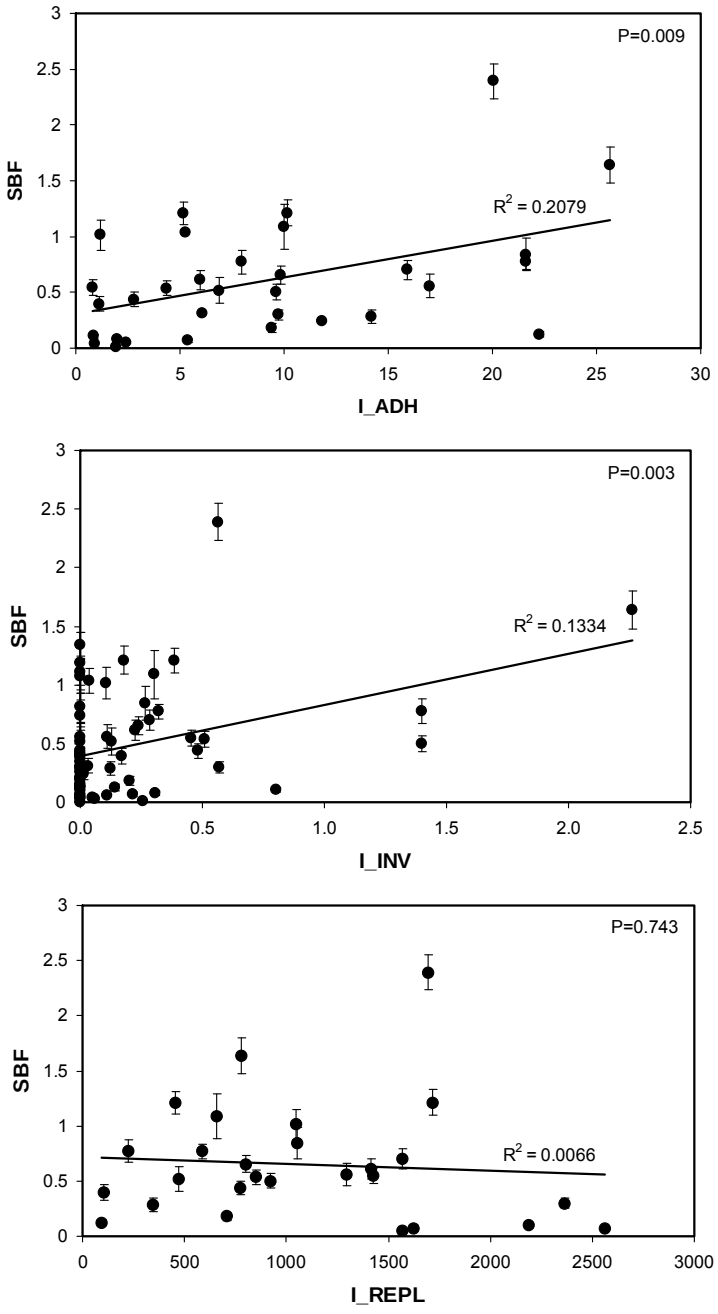


Figure 24. Specific biofilm formation (SBF) index of AIEC and non-AIEC strains isolated from intestinal mucosa. The mean SBF of AIEC strains was higher than for non-AIEC, as corroborated by one-way ANOVA (p=0.012).

Interestingly, higher adhesion indices from both AIEC and non-AIEC strains correlated with higher SBF indices (P=0.009). Moreover, the correlation was even stronger between the invasion and biofilm formation capacities of AIEC strains (P=0.003). No correlation was observed with the ability of AIEC strains to survive and replicate within macrophages (Figure 25).

Figure 25. Correlations between biofilm formation and the adhesion, invasion, and intra-macrophage replication abilities of both AIEC and non-AIEC strains.



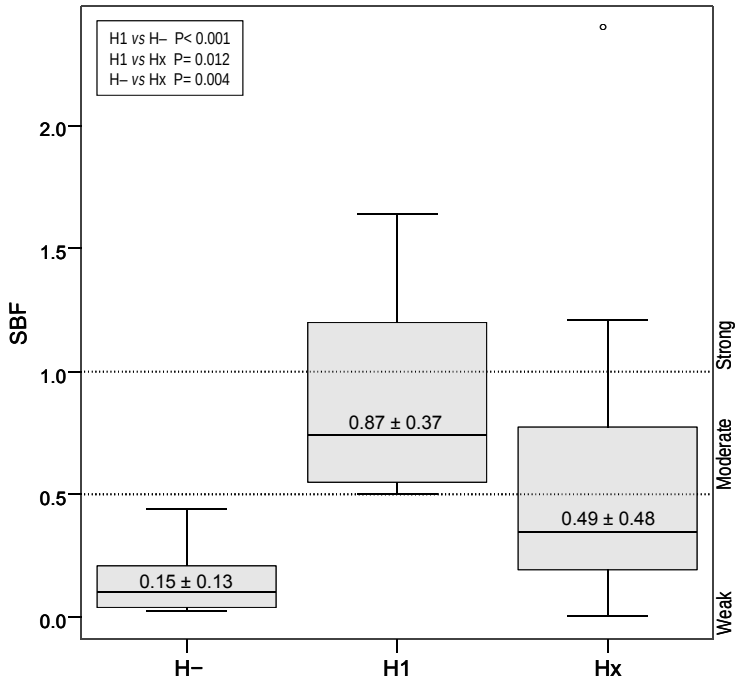
Adhesion and invasion indices correlated positively with biofilm formation capacity, whereas intra-macrophage survival and replication did not. Adhesion index was calculated as: I_ADH= attached bacterial cells/intestinal cell; invasion index as: I_INV(%) = (intracellular bacteria / 4×10^6 bacteria inoculated) × 100; and replication index as: I_REPL= (cfu ml^{-1} at 24h / cfu ml^{-1} at 1h)× 100.

Nonmotile strains were unable to form biofilms and, amongst motile strains, those with H1 flagellar type showed the highest biofilm formation indices

An additional factor that was associated with biofilm formation was the motility of the strains. Regardless of adhesion and invasion abilities, motile strains showed higher SBF indices than nonmotile strains (SBF_{MOTILE}= 0.61±0.48, $SBF_{NONMOTILE}$ = 0.14±0.13; P<0.001). All strains producing moderate-strong biofilms were motile, whereas strains classified as weak biofilm producers were heterogeneous in their motility capacities. In concordance, the isogenic mutant LF82-Δ*fliC* which is nonmotile, non-flagellated and express only few type 1 pili, did not display the ability to form biofilms (SBF= 0,393 ± 0,084) in contrast to LF82 wild type (SBF=1.641 ± 0.326).

Moreover, SBF indices were specifically higher for the H1 serotype as shown in Figure 26. All H1 serotypes were moderate-strong biofilm producers. In contrast, only 12 out of 33 (36.4%) of strains with other H types were classified within this category (Table 18).

Figure 26. Mean SBF index of motile and nonmotile strains irrespectively of their AIEC phenotype.



SBF indices were higher in motile strains, especially H1 serotypes, than nonmotile strains. H–: nonmotile strains; H1: motile and H1 flagellar type; Hx: motile and any flagellar type except for H1.

To determine whether motility and AIEC-like phenotype were intrinsically related factors, the frequency of motile and nonmotile strains within AIEC and

non-AIEC strains was calculated. Although the majority of AIEC strains were motile (81.5%), no significant differences were observed in comparison to non-AIEC strains (65.8%). Moreover, no interaction among these factors was detected by applying a factorial ANOVA. Therefore, motility and adherence/invasion capacity were independent factors associated with biofilm formation.

Table 18. Frequency of strains according to their mobility capacity and flagellar antigen type within biofilm producers and non-producers.

Motility/H type	Biofilm formation category				P
	Moderate-Strong (N= 26)		Weak (N= 39)		
	N	(%)	N	(%)	
H– (N=18)	0	(0)	18	(46.2)	< 0.001 [a] , 0.003 [b]
H1 (N=14)	14	(53.8)	0	(0)	< 0.001 [a] , < 0.001 [c]
Hx (N=33)	12	(46.2)	21	(53.8)	< 0.001 [c] , 0.003 [b]

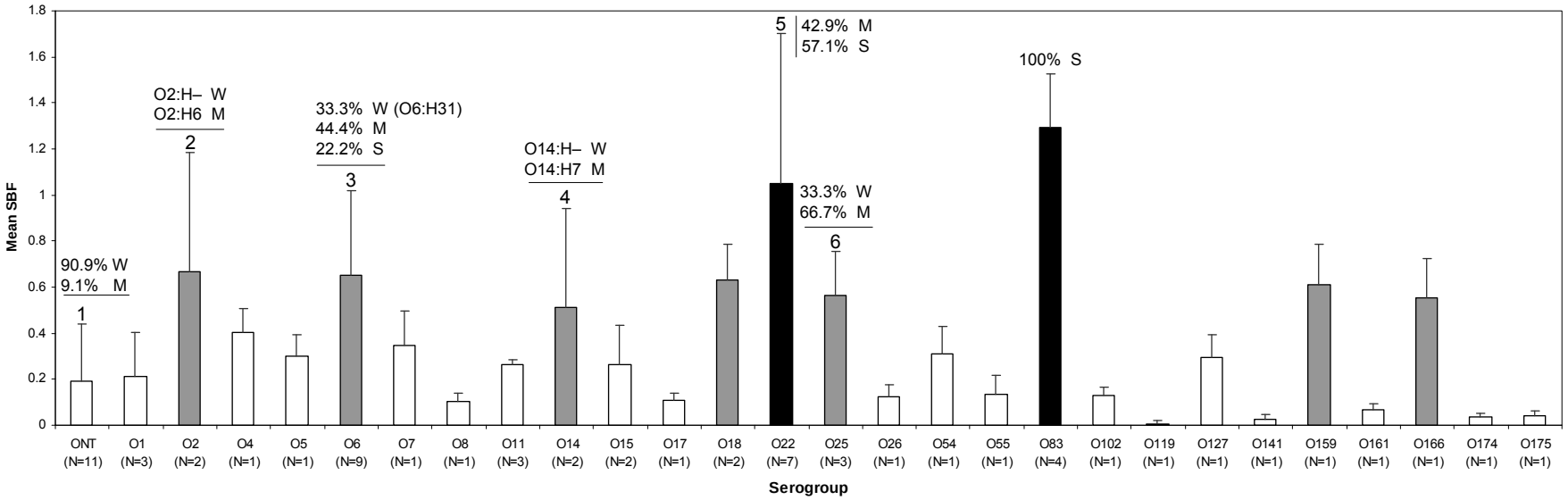
Abbreviators: H–: nonmotile strains; H1: motile and H1 flagellar type; Hx: motile and any flagellar type except H1.[a]significance between H– and H1; [b]significance between H– and Hx; [c]significance between H1 and Hx.

Serogroups associated with higher biofilm producing abilities

As shown in Figure 27, O83, followed by O22, showed the highest mean SBF indices. Regardless the AIEC phenotype and origin of the strains (intestinal or extraintestinal and non-IBD or CD associated), all the strains of O22 and O83 serogroup were found to be moderate-strong biofilm producers.

Other serogroups with mean SBF that fell into the 'moderate' category were: O2, O6, O14, O18, O25, O159, and O166. However, some strains that were unable to form biofilms were detected amongst these serogroups. For some serogroups such as O2 and O14 those strains classified as weak biofilm producers were particularly those nonmotile O2/O14 strains. In turn, strains with weak to strong biofilm formation abilities amongst the O6 strains belonged to a certain serotype (O6:H31) and this serotype was not present amongst the categories 'moderate' or 'strong' biofilm producers. Nevertheless, very few strains have been analyzed for some of these serogroups (O2, O14, O18, O25, O159, and O166) due to the nature of the strains isolated from the intestinal mucosa, thus no robust conclusions can be extracted for them.

Figure 27. Mean SBF index of the strains classified by their serogroup.



White bars: Serogroups with mean SBF that falls into 'weak' biofilm formation category. Grey bars: Serogroups with mean SBF that falls into 'moderate' biofilm formation category. Black bars: Serogroups with mean SBF that falls into 'strong' biofilm formation category. The serotype of those *E. coli* strains that showed different biofilm formation category than the mean SBF for the serogroup is specified:

1: Only AIEC17 (ONT:HNT) strain was classified as 'moderate' biofilm producer (M).

2: Nonmotile ECG-041 (O2:H–) strain was classified as 'weak' biofilm producer (W).

3: Three strains with O6:H31 serotype were classified as 'weak' biofilm producers, whereas strains with O6:H1, O6:H5 and O6:HNT serotypes were 'moderate' or 'strong' biofilm producers.

4: Nonmotile ECG-054 (O14:H–) was 'weak' biofilm producer (W).

5: Three strains were 'moderate' (O22:H1) and 4 strains 'strong' (O22:H1, O22:H7, and O22:H18) biofilm producers.

6: AIEC08 (O25:H4) was classified as 'weak' biofilm producer.

Distribution of virulence-associated genes and phylogroups within biofilm producers

Of the 65 *E. coli* strains used in this study, 45 (69.2%) harboured more than two virulence-associated genes in addition to *fimH*; thus, these strains are considered an extraintestinal pathogenic *E. coli* according to the definition of Johnson *et al* [32]. Virulence-associated gene distribution was similar between biofilm producers (moderate-strong) and non-biofilm producers (weak), with the exception of adherence factor *sfa/focDE* (S or F1C fimbriae) and the invasion-associated gene *ibeA* (Table 19), which were more prevalent in biofilm-forming strains (P=0.003 and P=0.017, respectively).

Although the *E. coli* collection studied was mainly composed of B2 (52.3%) and D (20%) phylogroups, significant differences were observed between the two categories of biofilm producers. As shown in Table 19, the B2 phylogroup was more frequent in moderate-strong biofilm forming strains (80.8% vs. 34.2%; P<0.001), whereas A and D phylogroups were more frequent within weak biofilm producers.

Table 19. Comparison of virulence gene prevalence and phylogroup between weak and moderate-strong biofilm producers.

			Biofilm formation category				
Virulence gene	**Total (N=65)**		Moderate-Strong (N=26)		Weak (N=39)		P
	N	(%)	N	(%)	N	(%)	
Adhesin-encoding genes							
papC	32	(49.2)	11	(42.3)	21	(53.8)	0.255
sfa/focDE	13	(20.0)	10	(38.5)	3	(7.7)	0.003
afa/draBC	8	(12.3)	2	(7.7)	6	(15.4)	0.301
fimH	62	(95.4)	26	(100)	36	(92.3)	0.209
fimAv$_{MT78}$	14	(21.5)	6	(23.1)	8	(20.5)	0.520
Protectin/invasion-encoding genes							
ibeA	9	(13.8)	7	(26.9)	2	(5.1)	0.017
K1 *neuC*	9	(13.8)	3	(11.5)	6	(15.4)	0.478
Siderophore-related genes							
iucD	37	(56.9)	13	(50.0)	24	(61.5)	0.253
Toxin-encoding genes							
hlyA	15	(23.1)	9	(34.6)	6	(15.4)	0.067
cnf1	15	(23.1)	9	(34.6)	6	(15.4)	0.067
cdtB	5	(7.7)	3	(11.5)	2	(5.1)	0.312
Phylogroup							
A	9	(13.8)	1	(3.8)	8	(21.1)	0.052
B1	8	(12.3)	3	(11.5)	5	(13.2)	0.583
B2	34	(52.3)	21	(80.8)	13	(34.2)	<0.001
D	13	(20.0)	1	(3.8)	12	(31.6)	0.006

Discussion

In this work, we describe the biofilm formation capacity of a recently described pathovar, adherent-invasive *E. coli* (AIEC), which is associated with Crohn's disease. The main result was that AIEC strains have stronger biofilm formation abilities than other *E. coli* strains isolated from the intestinal mucosa (non-AIEC). The latter share genotypic and phenotypic traits with AIEC [15] but lack the properties described for the pathovar: (i) adhesion to and invasion of intestinal epithelial cells *in vitro*, (ii) survival and replication capacity within macrophages without causing host-cell death, and (iii) induction of TNF-α release [11]. We also analyzed the relationship between biofilm formation, AIEC phenotype, serotype, and phylogroup, and the presence of virulence-associated genes.

As observed by other authors [33, 34], motility was a crucial factor for biofilm formation because none of the nonmotile strains were able to form biofilms (Table 18). This observation was further supported by the experiments performed with the isogenic mutant LF82-Δ*fliC*. Moreover, all 14 strains with H1 flagellar antigen were moderate-strong biofilm producers, in contrast to 46.2% of motile non-H1 types. Therefore, H1 flagellar antigen conferred, either directly or indirectly, an advantageous trait to form biofilms. Although motility was a necessary requirement for biofilm formation, it was not sufficient; 21 out of 47 motile strains were weak biofilm producers, indicating that additional factors are needed. In addition, strains with O2, O6, O14, O18, O22, O25, O83, O159 and O166 serogroups were found amongst the biofilm producers, in accordance with previous studies [25, 35]. Interestingly, the highest mean SBFs index was achieved by four strains that belonged to the O83 serogroup, in particular the O83:H1 serotype, being all the strains classified as strong biofilm producers. This group included two AIEC strains (AIEC reference strain LF82 [11], and the sepsis-associated strain PP16) and two non-AIEC strains (ECG-009 (isolated from two different CD patients) and ECG-043 (isolated from one non-IBD control) [15].

Some associations between biofilm-formation potential and some virulence-associated genes have been already described [24, 35-41]. In agreement with previous studies [25], the adhesin-coding gene *sfa/foc*DE was more frequently detected amongst biofilm producers. In addition, the gene *ibeA*, required for invasion in meningitis/sepsis-associated *E. coli* (MNEC) [42, 43], was more prevalent amongst strong biofilm producers. Interestingly, *ibeA*, in conjunction with *fimH* and $fimAv_{MT78}$, are virulence factors present in AIEC strain LF82 [16, 44].

Phylogenetic analyses have shown that *E. coli* strains fall into four main phylogenetic groups (A, B1, B2, and D) and that virulent ExPEC strains mainly belong to group B2 and, to a lesser extent, group D, whereas most commensal strains belong to group A [30, 42]. Although B2 was the most abundant phylogenetic group within the *E. coli* collection, B2 strains were significantly more prevalent amongst moderate-strong biofilm producers than weak biofilm producers ($P<0.001$), which were enriched in strains belonging to A and D phylogroups ($P=0.052$ and $P=0.006$ respectively). Of note, B2+D phylogroups are also more prevalent amongst *E. coli* strains from patients with CD or ulcerative colitis than in non-IBD controls [45].

The positive correlation between the levels of adhesion and invasion and the higher SBF indices lead to postulate that the machinery implicated to achieve the "AIEC phenotype" could share some factors necessary for biofilm formation, such as type 1 pili and flagella. Another possibility is that both processes could be related to a coordinated expression, for instance, by the EnvZ/OmpR regulatory system. Rohlion *et al* [46] recently proposed a model in which OmpC, a porin regulated by EnvZ/OmpR, has been implicated in the adherence-invasiveness of AIEC, and this system is also known to play an important role in biofilm formation [47]. The biofilm formation could also be dependent on the cyclic di-GMP concentration which was recently reported to regulate the expression of type 1 pili and flagella in AIEC reference strain LF82 [48].

Biofilms in the human gut are thought to play an agonistic role with the host [18], being necessary to achieve an homeostatic situation and appropriate gut physiology. Nevertheless, previous studies have highlighted the increased biofilm formation in patients with CD with respect to control subjects [49]. Moreover, the composition of the mucosa-associated microbiota is altered with respect to that of non-IBD controls [50]. It is widely accepted that the intestinal microbiota is essential to elicit the inflammation; however, the specific role of intestinal biofilms in CD is still uncertain. Changes in the composition and abundance of mucosa-associated biofilms have been proposed either to play a role in the onset or perpetuation of CD [49, 51-53] or to be a consequence of the defective immune regulation in CD patients [18, 54, 55]. Because we have analyzed the biofilm formation capacity of a collection of AIEC and non-AIEC strains using an *in vitro* method we can deduce that the ability of AIEC to form biofilms is irrespective of host factors. However, *in vivo* experiments would give interesting insights into the pathogenesis of AIEC in CD. Biofilm formation of AIEC in human gut, if confirmed, would confer to the pathovar an advantage for colonization of the intestine. Consequently, given the pathogenic behavior of

AIEC, a more stable colonization would increase their probability of invading the intestinal epithelium and further trigger mucosal inflammation and, possibly, granuloma formation. In this sense, and speculatively, biofilm formation could contribute to AIEC pathogenesis.

In conclusion, a novel phenotypic trait of AIEC pathovar was described in this work. Biofilm production ability of AIEC strains could be an additional trait involved in their pathogenesis. Further investigations to detect AIEC specific genetic determinants involved in biofilm formation and to analyze the genetic regulatory processes are essential to fully understand AIEC pathogenesis and elucidate a possible role of AIEC in CD.

References

1. Economou M, Pappas G: New Global Map of Crohn's Disease: Genetic, Environmental, and Socioeconomic Correlations. *Inflamm Bowel Dis* 2008, 14(5):709-720.

2. Baumgart DC, Carding SR: Inflammatory bowel disease: cause and immunobiology. *Lancet* 2007, 369(9573):1627-1640.

3. Xavier RJ, Podolsky DK: Unravelling the pathogenesis of inflammatory bowel disease. *Nature* 2007, 448(7152):427-434.

4. Halfvarson J, Bodin L, Tysk C, Lindberg E, Järnerot G: Inflammatory bowel disease in a Swedish twin cohort: a long-term follow-up of concordance and clinical characteristics. *Gastroenterology* 2003, 124(7):1767-1773.

5. De Hertogh G, Aerssens J, Geboes K, Geboes K: Evidence for the involvement of infectious agents in the pathogenesis of Crohn's disease. *World J Gastroenterol* 2008 14(6):845-852.

6. Hanauer S: Inflammatory Bowel Disease: Epidemiology, Pathogenesis, and Therapeutic Opportunities. *Inflamm Bowel Dis* 2006, 12:S3-S9.

7. Rutgeerts P, Goboes K, Peeters M, Hiele M, Penninckx F, Aerts R, Kerremans R, Vantrappen G: Effect of faecal stream diversion on recurrence of Crohn's disease in the neoterminal ileum. *Lancet* 1991, 338(8770):771-774.

8. Rutgeerts P, Hiele M, Geboes K, Peeters M, Penninckx F, Aerts R, Kerremans R: Controlled trial of metronidazole treatment for prevention of crohn's recurrence after ileal resection. *Gastroenterology* 1995, 108(6):1617-1621.

9. Sartor RB: Microbial Influences in Inflammatory Bowel Diseases. *Gastroenterology* 2008, 134(2):577-594.

10. Sellon RK, Tonkonogy S, Schultz M, Dieleman LA, Grenther W, Balish E, Rennick DM, Sartor RB: Resident Enteric Bacteria Are Necessary for Development of Spontaneous Colitis and Immune System Activation in Interleukin-10-Deficient Mice. *Infect Immun* 1998, 66(11):5224-5231.

11. Darfeuille-Michaud A, Boudeau J, Bulois P, Neut C, Glasser A-L, Barnich N, Bringer M-A, Swidsinski A, Beaugerie L, Colombel J-F: High prevalence of Adherent-Invasive *Escherichia coli* Associated With Ileal Mucosa in Crohn's Disease. *Gastroenterology* 2004, 127(2):412-421.

12. Martin HM, Campbell BJ, Hart CA, Mpofu C, Nayar M, Singh R, Englyst H, Williams HF, Rhodes JM: Enhanced *Escherichia coli* adherence and invasion in Crohn's disease and colon cancer. *Gastroenterology* 2004, 127(1):80-93.

13. Baumgart M, Dogan B, Rishniw M, Weitzman G, Bosworth B, Yantiss R, Orsi RH, Wiedmann M, McDonough P, Kim SG, Berg D, Schukken Y, Scherl E, Simpson KW: Culture independent analysis of ileal mucosa reveals a selective increase in invasive *Escherichia coli* of novel phylogeny relative to depletion of Clostridiales in Crohn's disease involving the ileum. *ISME J* 2007, 1(5):403-418.

14. Sasaki M, Sitaraman SV, Babbin BA, Gerner-Smidt P, Ribot EM, Garrett N, Alpern JA, Akyildiz A, Theiss AL, Nusrat A, Klapproth J-MA: Invasive *Escherichia coli* are a feature of Crohn's disease. *Lab Invest* 2007, 87(10):1042-1054.

15. Martinez-Medina M, Aldeguer X, Lopez-Siles M, González-Huix F, López-Oliu C, Dahbi G, Blanco JE, Blanco J, Garcia-Gil LJ, Darfeuille-Michaud A: Molecular diversity of *Escherichia coli* in the human gut: new ecological evidence supporting the role of adherent-invasive *E. coli* (AIEC) in Crohn's disease. *Inflamm Bowel Dis* 2009, 15(6):872-882.

16. Simpson KW, Dogan B, Rishniw M, Goldstein RE, Klaessig S, McDonough PL, German AJ, Yates RM, Russell DG, Johnson SE, Berg DE, Harel J, Bruant G, McDonough SP, Schukken YH: Adherent and Invasive *Escherichia coli* Is Associated with Granulomatous Colitis in Boxer Dogs. *Infect Immun* 2006, 74(8):4778-4792.

17. Hall-Stoodley L, Stoodley P: Biofilm formation and dispersal and the transmission of human pathogens. *Trends Microbiol* 2005, 13(1):7-10.

18. Everett ML, Palestrant D, Miller SE, Bollinger RR, Parker W: Immune exclusion and immune inclusion: a new model of host-bacterial interactions in the gut. *Clinical Applied Imm Rev* 2004, 4(5):321-332.

19. Anderson GG, Palermo JJ, Schilling JD, Roth R, Heuser J, Hultgren SJ: Intracellular Bacterial Biofilm-Like Pods in Urinary Tract Infections. *Science* 2003, 301(5629):105-107.

20. Kaper JB, Nataro JP, Mobley HLT: Pathogenic *Escherichia coli*. *Nat Rev Microbiol* 2004, 2(2):123-140.

21. Boudeau J, Glasser A-L, Masseret E, Joly B, Darfeuille-Michaud A: Invasive Ability of an *Escherichia coli* Strain Isolated from the Ileal Mucosa of a Patient with Crohn's Disease. *Infect Immun* 1999, 67(9):4499-4509.

22. Blanco M, Blanco JE, Alonso MP, Mora A, Balsalobre C, Munoa F, Juárez A, Blanco J: Detection of pap, sfa and afa adhesin-encoding operons in uropathogenic *Escherichia coli* strains: Relationship with expression of adhesins and production of toxins. *Res Microbiol* 1997, 148(9):745-755.

23. Nicolas-Chanoine M-H, Blanco J, Leflon-Guibout V, Demarty R, Alonso MP, Canica MM, Park Y-J, Lavigne J-P, Pitout J, Johnson JR: Intercontinental emergence of *Escherichia coli* clone O25:H4-ST131 producing CTX-M-15. *J Antimicrob Chemother* 2008, 61(2):273-281.

24. Danese P, Pratt L, Dove S, Kolter R: The outer membrane protein, Antigen 43, mediates cell-to-cell interactions within *Escherichia coli* biofilms. *Mol Microbiol* 2000, 37(2):424-432.

25. Naves P, del Prado G, Huelves L, Gracia M, Ruiz V, Blanco J, Dahbi G, Blanco M, del Carmen Ponte M, Soriano F: Correlation between virulence factors and in vitro biofilm formation by *Escherichia coli* strains. *Microb Pathog* 2008, 45(2):86-91.

26. Naves P, Prado Gd, Huelves L, Gracia M, Ruiz V, J. Blanco, Rodríguez-Cerrato V, Ponte MC, Soriano F: Measurement of biofilm formation by clinical isolates of *Escherichia coli* is method-dependent. *J Appl Microbiol* 2008, 105(2):585-590.

27. Niu C, Gilbert ES: Colorimetric Method for Identifying Plant Essential Oil Components That Affect Biofilm Formation and Structure. *Appl Environ Microbiol* 2004, 70(12):6951-6956.

28. Glasser A-L, Boudeau J, Barnich N, Perruchot M-H, Colombel J-F, Darfeuille-Michaud A: Adherent Invasive *Escherichia coli* Strains from Patients with Crohn's Disease Survive and Replicate within Macrophages without Inducing Host Cell Death. *Infect Immun* 2001, 69(9):5529-5537.

29. Guinée PA, Agterberg CM, Jansen WH: *Escherichia coli* O antigen typing by means of a mechanized microtechnique. *Appl Microbiol* 1972, 24(1):127-131.

30. Clermont O, Bonacorsi S, Bingen E: Rapid and Simple Determination of the *Escherichia coli* Phylogenetic Group. *Appl Environ Microbiol* 2000, 66(10):4555-4558.

31. Blanco M, Blanco J, Dahbi G, Alonso M, Mora A, Coira M, Madrid C, Juárez A, Bernárdez M, González E, Blanco J: Identification of two new intimin types in atypical enteropathogenic *Escherichia coli. Int Microbiol* 2006, 9(2):103-110.

32. Johnson JR, Murray AC, Gajewski A, Sullivan M, Snippes P, Kuskowski MA, Smith KE: Isolation and Molecular Characterization of Nalidixic Acid-Resistant Extraintestinal Pathogenic *Escherichia coli* from Retail Chicken Products. *Antimicrob Agents Chemother* 2003, 47(7):2161-2168.

33. Pratt LA, Kolter R: Genetic analysis of *Escherichia coli* biofilm formation: roles of flagella, motility, chemotaxis and type I pili. *Mol Microbiol* 1998, 30(2):285-293.

34. Van Houdt R, Michiels CW: Role of bacterial cell surface structures in *Escherichia coli* biofilm formation. *Res Microbiol* 2005, 156(5-6):626-633.

35. Kanamaru S, Kurazono H, Terai A, Monden K, Kumon H, Mizunoe Y, Ogawa O, Yamamoto S: Increased biofilm formation in *Escherichia coli* isolated from acute prostatitis. *Int J Antimicrob Agents* 2006, 28(Supplement 1):21-25.

36. Ong C-LY, Ulett GC, Mabbett AN, Beatson SA, Webb RI, Monaghan W, Nimmo GR, Looke DF, McEwan AG, Schembri MA: Identification of Type 3 Fimbriae in Uropathogenic *Escherichia coli* Reveals a Role in Biofilm Formation. *J Bacteriol* 2008, 190(3):1054-1063.

37. Schembri MA, Dalsgaard D, Klemm P: Capsule Shields the Function of Short Bacterial Adhesins. *J Bacteriol* 2004, 186(5):1249-1257.

38. Soto SM, Smithson A, Martinez JA, Horcajada JP, Mensa J, Vila J: Biofilm Formation in Uropathogenic *Escherichia coli* Strains: Relationship With Prostatitis, Urovirulence Factors and Antimicrobial Resistance. *J Urol* 2007, 177(1):365-368.

39. Ulett GC, Mabbett AN, Fung KC, Webb RI, Schembri MA: The role of F9 fimbriae of uropathogenic *Escherichia coli* in biofilm formation. *Microbiology* 2007, 153(7):2321-2331.

40. Ulett GC, Valle J, Beloin C, Sherlock O, Ghigo J-M, Schembri MA: Functional Analysis of Antigen 43 in Uropathogenic *Escherichia coli* Reveals a Role in Long-Term Persistence in the Urinary Tract. *Infect Immun* 2007, 75(7):3233-3244.

41. Vianney A, Jubelin G, Renault S, Dorel C, Lejeune P, Lazzaroni JC: *Escherichia coli tol* and *rcs* genes participate in the complex network affecting curli synthesis. *Microbiology* 2005, 151(7):2487-2497.

42. Bidet P, Mahjoub-Messai F, Blanco J, Blanco J, Dehem M, Aujard Y, Bingen E, Bonacorsi S: Combined Multilocus Sequence Typing and O Serogrouping Distinguishes *Escherichia coli* Subtypes Associated with Infant Urosepsis and/or Meningitis. *J Infect Dis* 2007, 196(2):297-303.

43. Xie Y, Kim KJ, Kim KS: Current concepts on *Escherichia coli* K1 translocation of the blood-brain barrier. *FEMS Immunol Med Microbiol* 2004, 42(3):271-279.

44. Boudeau J, Barnich N, Darfeuille-Michaud A: Type 1 pili-mediated adherence of *Escherichia coli* strain LF82 isolated from Crohn's disease is involved in bacterial invasion of intestinal epithelial cells. *Mol Microbiol* 2001, 39(5):1272-1284.

45. Kotlowski R, Bernstein CN, Sepehri S, Krause DO: High prevalence of *Escherichia coli* belonging to the B2+D phylogenetic group in inflammatory bowel disease. *Gut* 2007, 56(5):669-675.

46. Rolhion N, Carvalho FA, Darfeuille-Michaud A: OmpC and the sigma(E) regulatory pathway are involved in adhesion and invasion of the Crohn's disease-associated *Escherichia coli* strain LF82. *Mol Microbiol* 2007, 63(6):1684-1700.

47. Pruss BM, Besemann C, Denton A, Wolfe AJ: A Complex Transcription Network Controls the Early Stages of Biofilm Development by *Escherichia coli*. *J Bacteriol* 2006, 188(11):3731-3739.

48. Claret L, Miquel S, Vieille N, Ryjenkov DA, Gomelsky M, Darfeuille-Michaud A: The Flagellar Sigma Factor FliA Regulates Adhesion and Invasion of Crohn Disease-associated *Escherichia coli* via a Cyclic Dimeric GMP-dependent Pathway. *J Biol Chem* 2007, 282(46):33275-33283.

49. Swidsinski A, Ladhoff A, Pernthaler A, Swidsinski S, Loening-Baucke V, Ortner M, Weber J, Hoffmann U, Schreiber S, Dietel M, Lochs H: Mucosal flora in inflammatory bowel disease. *Gastroenterology* 2002, 122(1):44-54.

50. Martinez-Medina M, Aldeguer X, Gonzalez-Huix F, Acero D, Garcia-Gil LJ: Abnormal microbiota composition in the ileocolonic mucosa of Crohn's disease patients as revealed by polymerase chain reaction-denaturing gradient gel electrophoresis. *Inflamm Bowel Dis* 2006, 12(12):1136-1145.

51. Dicksved J, Halfvarson J, Rosenquist M, Jarnerot G, Tysk C, Apajalahti J, Engstrand L, Jansson JK: Molecular analysis of the gut microbiota of identical twins with Crohn's disease. *ISME J* 2008, 2(7):716-727.

52. Kleessen B, Kroesen A, Buhr H, Blaut M: Mucosal and invading bacteria in patients with inflammatory bowel disease compared with controls. *Scand J Gastroenterol* 2002, 37(9):1034-1041.

53. Schultsz C, van den Berg FM, ten Kate FW, Tytgat GNJ, Dankert J: The intestinal mucus layer from patients with inflammatory bowel disease harbors high numbers of bacteria compared with controls. *Gastroenterology* 1999, 117(5):1089-1097.

54. Lupp C, Robertson ML, Wickham ME, Sekirov I, Champion OL, Gaynor EC, Finlay BB: Host-mediated inflammation disrupts the intestinal microbiota and promotes the overgrowth of *Enterobacteriaceae*. *Cell Host Microbe* 2007, 2(2):119-129.

55. Wehkamp J, Stange EF: Is there a role for defensins in IBD? *Inflamm Bow Dis* 2008, 14(S2):S85-S87.

3.2

Similarity and divergence among adherent-invasive *E. coli* (AIEC) and extraintestinal pathogenic *E. coli* (ExPEC)

Introduction

Enterobacteriaceae, especially *Escherichia coli*, have been repeatedly suggested to play a role in the origin and/or perpetuation of Crohn's disease (CD). In part, this suggestion was based on the higher abundance of this bacterium in CD patients than in control subjects [1-9]. Although considerable effort has been devoted to the search for intestinal pathogenic *E. coli* associated with CD, to date none of the six previously described pathovars [10] has been implicated in this condition. Darfeuille-Michaud *et al.* [2] observed that *E. coli* strains with adhesion and invasion properties colonised the ileal mucosa of CD patients more frequently than that of control subjects. Darfeuille-Michaud *et al.* further characterised these strains and proposed a new potentially *E. coli* pathovar associated with CD, which was designated adherent-invasive *E. coli* (AIEC) [11]. The implication of AIEC in CD is achieving increasing relevance because several independent studies from different countries have reported a higher prevalence of invasive *E. coli* in CD patients [1, 5, 12-14].

The main characteristics of AIEC are: 1) the ability to adhere to and invade intestinal epithelial cells, 2) the ability to survive and replicate expansively within macrophages without triggering host cell death, and inducing the release of TNFα [15], and 3) the lack of known invasive determinants [12]. Recently, Glasser *et al.* [16] proposed a model explaining the mechanism of pathogenesis for AIEC strains. The AIEC strains isolated to date are clonally diverse and belong to distinct serotypes. Moreover, despite the fact that they primarily fall into the B2 phylogroup, AIEC strains belonging to the A, B1, and D phylogroups have also been isolated [1, 5, 13, 14, 17]. Although no specific virulence factors have been described for this pathovar, AIEC strains carry many virulence-associated genes characteristic of extraintestinal pathogenic *E. coli* (ExPEC) strains, which suggests that the AIEC pathovar could be closely related to the ExPEC pathovar [1, 12, 13].

The aim of this work was to determine the frequency of strains with 'AIEC phenotype' amongst *E. coli* strains that cause extraintestinal infections, including uropathogenic *E. coli* (UPEC), septicemic *E. coli*, and neonatal meningitis *E. coli*

(NMEC). To achieve this objective, we determined the ability of a collection of ExPEC strains to adhere to and invade intestinal epithelial cells, as well as their capacity to survive and replicate within macrophages. In parallel, we compared the distribution of virulence-associated genes amongst ExPEC and AIEC strains. Furthermore, we searched for a common phylogenetic origin of the ExPEC strains that had an AIEC phenotype (referred to in this study as extraintestinal AIEC) and a collection of AIEC strains, mainly isolated from the intestinal mucosa of CD patients (intestinal AIEC).

Materials and methods

Bacterial strains

The present study involved a collection of 86 *E. coli* strains, some of which were the same as those reported in previously published studies [12, 13, 18-23] (Table 20). Sixty-three (73.3%) were obtained from human extraintestinal infections (28 from urinary tract infections, 21 from sepsis, 12 from meningitis, 1 from intra-abdominal pus, and 1 from a wound infection) and 23 were obtained from the intestinal mucosa of patients with CD (16 strains) or ulcerative colitis (UC) (1 strain) and the intestinal mucosa of control subjects (non-IBD) (6 strains). Control subjects were asymptomatic and did not present inflammation and/or evidence of polyps during colonoscopy. Among CD patients, 39% had Crohn's colitis (C-CD), 35% had Crohn's ileitis (I-CD), and 26% had ileal/colonic disease (IC-CD). Further information about the source of intestinal AIEC strains can be obtained from reference [13]. The prototype AIEC strain LF82 was included in this group of 23 intestinal AIEC strains.

Adhesion and invasion assays with Intestine-407 epithelial cells

Adhesion (I_ADH) and invasion indices (I_INV) were determined as explained in chapter 2 (quantitative analysis) following the gentamicin protection assay previously described elsewhere [11].

Survival and replication in macrophages J774

Intra-macrophage replication index (I_REPL) was performed as explained in chapter 2, following the method previously described elsewhere [12, 15].

Those strains with I_INV > 0.1 and I_REPL > 100% were classified as AIEC strains in the present study.

Phylotyping and virulence genotyping by PCR

Determination of the major *E. coli* phylogenetic group (A, B1, B2, or D) was performed as described by Clermont *et al.* [24].

The presence of virulence genes was analysed as described elsewhere [13]. Primers specific for 10 genes and operons that encode extraintestinal virulence factors characteristic of ExPEC were used. These genes included adhesins (pyelonephritis-associated pili [*papC*], S and F1C fimbriae [*sfa/focDE*], Dr-binding adhesins [*afa/draBC*], and type 1 fimbriae [*fimH* and $fimAv_{MT78}$, the avian pathogenic variant of *fimA*]), two toxins (*hlyA* and *cnf1*), and one aerobactin (*iucD*). The analysed genes also included two protectin/invasion-encoding genes corresponding to the K1 *kps* variant (*neuC*) and the invasion of brain endothelium gene (*ibeA*). The *papC*-positive strains were tested for the *papG* I, *papG* II, and *papG* III alleles. The *E. coli* were also screened for specific genes found in diarrhoeagenic *E. coli* pathovars (*stx1*, *stx2*, *eae*, *bfpA*, *ipaH*, *pCDV432*, *eltA,* and *est*).

Additional virulence genes (*cdtB*, cytolethal distending toxin*; bmaE*, M fimbriae; *gafD*, G fimbriae; *sat*, secreted autotransporter toxin; *cvaC*, microcin (colicin) V; *traT*, serum-resistance associated; *malX*, pathogenicity island marker; *usp*, uropathogenic specific protein; *focG*, F1C fimbriae; *sfaS*, S fimbriae; *iroN*, salmochelin receptor; *kpsM-II*, group 2 capsule, and *kpsM-III*, group 3 capsule) were investigated in those strains included in Figures 29 and 30. The amplification procedures have been documented elsewhere ([25] and references therein).

Serotyping

Determination of O and H antigens was carried out using the method previously described by Guinée *et al.* [26].

Pulsed-field gel electrophoresis (PFGE)

PFGE was performed as described elsewhere [27]. Agarose-embedded DNA was digested with 0.2 U/µl *Xba*I (Roche) according to manufacturer's instructions. The *Xba*I-digested genomic DNA was analysed by 1% agarose gel in 0.5x Tris-boric acid-EDTA (TBE) buffer at 14ºC using CHEF MAPPER (BioRad). The gel was run for 21.30 h at 6 V/cm, with initial and final switch times of 2.16 s and 54.17 s, respectively. The gel was stained with ethidium bromide (1µg/mL), observed using a Gel Doc 2000 system (BioRad), and analysed with the BioNumerics fingerprinting software (Applied Maths, St-Martens-Latem, Belgium). Cluster analysis of the Dice similarity indices based on the unweighted

pair group method using arithmetic averages (UPGMA) was performed to generate a dendrogram describing the relationships amongst the PFGE profiles.

Multilocus sequence typing (MLST)

Multilocus sequence typing (MLST) was carried out as previously described [28]. Gene amplification and sequencing of the seven housekeeping genes (*adk*, *fumC*, *gyrB*, *icd*, *mdh*, *purA*, and *recA*) were performed using the primers and protocol specified at the *E. coli* MLST website (http://mlst.ucc.ie/mlst/dbs/Ecoli). The sequences were reviewed by visual inspection with the BioEdit Sequence Alignment Editor (version 7.0.9; Ibis Biosciences). The ClustalW2 programme was used to align the sequences. The allelic profile of the seven gene sequences, the sequence types (STs), and the sequence complexes (Cplx, defined as STs that are linked by distances of one or two allelic differences) were obtained via the electronic database at the *E. coli* MLST website.

Table 20. Collection of ExPEC and intestinal AIEC strains used in this study. For each strain, the adhesion (I_ADH), invasion (I_INV), and intra-macrophage replication (I_REPL) abilities are expressed as the mean and standard deviation of triplicate assays. Seropathotype, phylogenetic affiliation, and type of extended-spectrum β-lactamase are also specified. References indicate the origin of the strain isolation.

[1]For those strains of intestinal origin, specific zones from which the strains have been isolated along the intestinal tract are indicated in parenthesis.

[2]Calculated as the mean number of bacteria per I407 cell after 3 hours of incubation.

[3]Calculated as the percentage of inoculum surviving after 1 hour of gentamicin treatment: I_INV (%) = (intracellular bacteria / 4×10^6 bacteria inoculated) × 100.

[4]Percentage of intracellular bacteria at 24 hours postinfection relative to the number after 1 hour of gentamicin treatment: I_REPL (%) = (cfu ml^{-1} at 24h / cfu ml^{-1} at 1h) × 100.

Abbreviators: EI: extraintestinal; I: intestinal; I-CD: Crohn's ileitis; IC-CD: ileocolonic disease; C-CD: Crohn's colitis; UC: ulcerative colitis; non-IBD: controls without inflammatory bowel disease; UTI: urinary tract infection; H–: non-motile strain; ExPEC-like: strains with ≥ 2 virulence-associated genes regardless of the presence of *fimH*.

Those strains with I_INV > 0.1 and I_REPL > 100% were classified as AIEC strains in the present study.

Table 20. Collection of ExPEC and intestinal AIEC strains used in this study.

STRAIN	ORIGIN[1]		Country	AIEC phenotype	I_ ADH[2]	I_INV[3]	I_REPL[4]	Serotype	Phylo.	ExPEC-like	β-lactamase	Virulence gene carriage	Reference
15802	EI	Intra-abdominal pus	Canada	−	2.2 ± 2.7	0.009 ± 0.005		O25:H4	B2	−	CTX-M-15	*iucD, fimH*	[23]
EC-1	EI	wound infection	Spain	−	1.2 ± 0.4	0.026 ± 0.009		O6:H1	B2	+	−	*papC, papGIII, sfa/focDE, hlyA, cnf1, fimH*	[22]
SM22	EI	meningitis	USA	−	0.2 ± 0.3	0.005 ± 0.006		O6:H1	B2	−	−	*iucD, fimH*	[18]
H1166	EI	meningitis	France	−	1.8 ± 0.8	0.060 ± 0.057		O6:H1	B2	+	−	*papC, papGII, sfa/focDE, iucD, hlyA, fimH*	[18]
SM18	EI	meningitis	USA	−	0.1 ± 0.1	0.000 ± 0.000		O7:H−	D	+	−	*papC, papGI-GII, neuC, iucD, fimH*	[18]
SM21	EI	meningitis	USA	−	0.8 ± 1.3	0.007 ± 0.003		O16:H6	B2	+	−	*papC, papGII, neuC, iucD, fimH, fimAv$_{MT78}$*	[18]
SM43	EI	meningitis	France	−	0.1 ± 0.1	0.013 ± 0.018		O6:H1	B2	+	−	*sfa/focDE, iucD, hlyA, cnf1, fimH*	[18]
SM57	EI	meningitis	USA	−	0.6 ± 0.2	0.013 ± 0.004		O83:H7	B2	+	−	*sfa/focDE, neuC, iucD, ibeA, fimH, fimAv$_{MT78}$*	[18]
SM63	EI	meningitis	USA	−	0.5 ± 0.1	0.024 ± 0.016		O1:H7	B2	+	−	*papC, papGI, neuC, iucD, fimH, fimAv$_{MT78}$*	[18]
SM69	EI	meningitis	France	−	1.8 ± 0.4	0.018 ± 0.011		O18:H7	B2	+	−	*papC, sfa/focDE, neuC, iucD, hlyA, cnf1, ibeA, fimH, fimAv$_{MT78}$*	[18]
SM72	EI	meningitis	France	−	0.2 ± 0.0	0.011 ± 0.005		O45:H7	B2	+	−	*papC, papGII, sfa/focDE, neuC, iucD, fimH, fimAv$_{MT78}$*	[18]
SM148	EI	meningitis	France	−	0.1 ± 0.1	0.009 ± 0.012		O18:H7	B2	+	−	*sfa/focDE, neuC, iucD, ibeA, fimH, fimAv$_{MT78}$*	[18]
SM168	EI	meningitis	France	−	0.8 ± 0.3	0.013 ± 0.011		O18:H7	B2	+	−	*sfa/focDE, neuC, iucD, hlyA, cnf1, ibeA, fimH, fimAv$_{MT78}$*	[18]
SM177	EI	meningitis	France	−	0.0 ± 0.0	0.000 ± 0.000		O1:H−	D	+	−	*papC, papGI-GII, neuC, iucD, hlyA*	[18]
H1088	EI	sepsis	Spain	−	1.8 ± 2.0	0.043 ± 0.053		O25:H4	B2	−	−	*iucD, fimH*	This study
H109	EI	sepsis	Spain	−	0.0 ± 0.0	0.002 ± 0.003		O6:H10	A	−	−	*iucD*	This study
H685	EI	sepsis	Spain	−	0.5 ± 0.1	0.021 ± 0.002		O25:H4	B2	−	−	*ibeA, fimH*	This study
H6166	EI	sepsis	France	−	0.7 ± 0.2	0.023 ± 0.025		O45:H7	B2	+	−	*papC, papGII, sfa/focDE, neuC, iucD, hlyA, cnf1, fimH, fimAv$_{MT78}$*	[18]
H6290	EI	sepsis	France	−	0.5 ± 0.1	0.018 ± 0.004		O45:H7	B2	+	−	*papC, neuC, iucD, fimH, fimAv$_{MT78}$*	[18]
FV7561	EI	sepsis	Spain	−	1.3 ± 0.5	0.041 ± 0.030		O25:H4	B2	+	CTX-M-15	*afa/draBC, iucD, fimH*	[23]
H169	EI	sepsis	Spain	−	1.5 ± 0.0	0.029 ± 0.009		O18:H7	B2	+	−	*sfa/focDE, neuC, iucD, ibeA, fimH, fimAv$_{MT78}$*	[21]
PP16	EI	sepsis	Spain	+	1.2 ± 1.0	0.104 ± 0.081	1045.9 ± 541.2	O83:H1	B2	+	−	*ibeA, fimH, fimAv$_{MT78}$*	This study
H102A	EI	sepsis	Spain	−	0.8 ± 0.3	0.031 ± 0.013		O83:H1	B2	+	−	*ibeA, fimH, fimAv$_{MT78}$*	This study
H126	EI	sepsis	Spain	−	0.3 ± 0.3	0.027 ± 0.011		O83:H31	B2	+	−	*iucD, ibeA, fimH*	This study
H106A	EI	sepsis	Spain	−	0.2 ± 0.2	0.001 ± 0.001		O6:H31	B2	+	−	*papC, papGIII, sfa/focDE, iucD, hlyA, cnf1, fimH*	This study
PP215	EI	sepsis	Spain	+	0.8 ± 0.6	0.453 ± 0.350	1425.4 ± 229.4	O6:H1	B2	+	−	*papC, papGII, sfa/focDE, iucD, hlyA, cnf1, fimH*	This study
H219B	EI	sepsis	Spain	−	0.2 ± 0.1	0.006 ± 0.009		O25:H4	B2	+	−	*iucD, ibeA, fimH, fimAv$_{MT78}$*	This study
H676A	EI	sepsis	Spain	−	0.1 ± 0.1	0.003 ± 0.004		O25:H4	B2	+	−	*papC, papGIII, iucD, hlyA, ibeA, fimH, fimAv$_{MT78}$*	This study
H68a	EI	sepsis	Spain	−	0.1 ± 0.1	0.005 ± 0.000		O6:H7	B2	+	−	*papC, papGIII, sfa/focDE, hlyA, cnf1, ibeA, fimH*	[19]
H778	EI	sepsis	Spain	−	1.7 ± 1.8	0.033 ± 0.039		O25:H4	B2	+	−	*afa/draBC, iucD, ibeA, fimH*	This study
H810A	EI	sepsis	Spain	−	0.3 ± 0.1	0.016 ± 0.013		O25:H4	B2	+	−	*iucD, ibeA, fimH*	This study
H858	EI	sepsis	Spain	−	3.6 ± 2.0	0.083 ± 0.018		O25:H4	B2	+	−	*afa/draBC, ibeA, fimH*	This study
PP209	EI	sepsis	Spain	−	0.6 ± 0.6	0.029 ± 0.002		O6:H1	B2	+	−	*papC, papGII-GIII, sfa/focDE, iucD, hlyA, cnf1, fimH*	This study
PP42	EI	sepsis	Spain	−	0.0 ± 0.1	0.019 ± 0.009		O25:H4	B2	−	−	*ibeA, fimH*	This study
PP89	EI	sepsis	Spain	−	0.0 ± 0.0	0.030 ± 0.014		O6:H10	A	−	−	*iucD*	This study
EC-2	EI	UTI	Spain	−	3.0 ± 0.1	0.038 ± 0.032		ONT:HNT	B1	−	−	*fimH*	[22]
VC1	EI	UTI	France	−	1.8 ± 1.9	0.035 ± 0.035		O25:H4	B2	−	−	*iucD, fimH*	This study
HDE3	EI	UTI	France	−	0.1 ± 0.0	0.004 ± 0.001		O25:H4	B2	−	CTX-M-15	*iucD, fimH*	[23]
17102	EI	UTI	Canada	−	5.2 ± 0.9	0.078 ± 0.043		O25:H4	B2	−	CTX-M-15	*iucD, fimH*	[23]
OL52A	EI	UTI	Spain	−	0.0 ± 0.0	0.004 ± 0.005		O101:H−	A	−	−	*iucD*	[20]
OL80A	EI	UTI	Spain	−	0.2 ± 0.2	0.008 ± 0.004		O51:H49	B2	−	−	*fimH, fimAv$_{MT78}$*	[20]
OB59A	EI	UTI	Spain	−	0.2 ± 0.1	0.011 ± 0.006		O83:H1	B2	−	−	*ibeA, fimH*	[20]

STRAIN	ORIGIN[1]		Country	AIEC phenotype	I_ ADH[2]		I_INV[3]		I_REPL[4]		Serotype	Phylo.	ExPEC-like	β-lactamase	Virulence gene carriage	Reference
COR227	EI	UTI	Spain	−	0.6	± 0.8	0.037	± 0.010			O6:H25	B1	−	−	*fimH*	[20]
FV7563	EI	UTI	Spain	+	6.9	± 5.9	0.129	± 0.072	470.0	± 264.0	O25:H4	B2	+	CTX-M-15	*afa/draBC, iucD, fimH*	[23]
FV7569	EI	UTI	Spain	−	6.2	± 2.5	0.083	± 0.054			O25:H4	B2	+	CTX-M-15	*afa/draBC, iucD, fimH*	[23]
FV7588	EI	UTI	Spain	−	2.7	± 1.0	0.060	± 0.028			O25:H4	B2	+	CTX-M-15	*afa/draBC, iucD, fimH*	[23]
OL96A	EI	UTI	Spain	+	5.2	± 5.0	0.388	± 0.159	457.5	± 259.3	O6:H1	B2	+	−	*papC, papGII, sfa/focDE, iucD, hlyA, cnf1, fimH*	[20]
OL61A	EI	UTI	Spain	−	0.5	± 0.2	0.026	± 0.014			O75:H7	B2	+	−	*sfa/focDE, neuC, hlyA, cnf1, ibeA, fimH*	[20]
OL37A	EI	UTI	Spain	−	0.1	± 0.1	0.000	± 0.000			O1:H1	D	+	−	*papC, neuC, iucD, fimH, fimAv$_{MT78}$*	[20]
OL65A	EI	UTI	Spain	−	1.4	± 1.2	0.003	± 0.004			O6:H1	B2	+	−	*sfa/focDE, iucD, hlyA, cnf1, fimH*	[20]
OB112A	EI	UTI	Spain	−	0.5	± 0.3	0.005	± 0.000			O2:H−	B2	+	−	*papC, papGII-GIII, sfa/focDE, iucD, hlyA, cnf1, fimH*	[20]
OB29A	EI	UTI	Spain	−	0.2	± 0.2	0.029	± 0.023			O6:H1	B2	+	−	*papC, papGIII, sfa/focDE, hlyA, cnf1, fimH*	[20]
OB4A	EI	UTI	Spain	−	0.8	± 0.2	0.017	± 0.009			O6:H1	B2	+	−	*sfa/focDE, hlyA, cnf1*	[20]
OL100A	EI	UTI	Spain	−	0.2	± 0.1	0.000	± 0.000			O18:H−	B2	+	−	*papC, papGII, sfa/focDE, iucD, hlyA, fimH*	[20]
OB64A	EI	UTI	Spain	−	0.3	± 0.3	0.002	± 0.003			O75:H5	B2	+	−	*papC, papGIII, sfa/focDE, hlyA, cnf1, ibeA, fimH*	[20]
OL16A	EI	UTI	Spain	−	0.4	± 0.4	0.000	± 0.000			O6:H−	B2	+	−	*papC, papGII, sfa/focDE, neuC, iucD, hlyA, cnf1, fimH*	[20]
OB103A	EI	UTI	Spain	−	1.2	± 0.8	0.016	± 0.008			O6:H1	B2	+	−	*sfa/focDE, hlyA, cnf1, fimH*	[20]
OL85A	EI	UTI	Spain	−	0.4	± 0.2	0.024	± 0.015			O2:H1	B2	+	−	*papC, papGIII, sfa/focDE, iucD, hlyA, cnf1, fimH*	[20]
OB23A	EI	UTI	Spain	−	1.1	± 0.1	0.072	± 0.018			O6:H1	B2	+	−	*papC, papGIII, sfa/focDE, hlyA, cnf1, fimH*	[20]
OB123A	EI	UTI	Spain	−	3.7	± 0.5	0.039	± 0.008			O6:H1	B2	+	−	*papC, papGIII, sfa/focDE, hlyA, cnf1, fimH*	[20]
OB79A	EI	UTI	Spain	−	0.5	± 0.2	0.014	± 0.008			O83:H1	B2	+	−	*ibeA, fimH, fimAv$_{MT78}$*	[20]
OL64A	EI	UTI	Spain	−	0.0	± 0.0	0.000	± 0.000			O6:H31	B2	+	−	*papC, papGIII, sfa/focDE, hlyA, cnf1, fimH*	[20]
OL118A	EI	UTI	Spain	−	0.2	± 0.1	0.003	± 0.004			O6:H31	B2	+	−	*papC, papGIII, sfa/focDE, hlyA, cnf1, fimH*	[20]
AIEC25	I	C-CD (colon)	Spain	+	2.8	± 1.3	0.482	± 0.129	775.93	± 128.3	O6:H31	B2	+	−	*papC, papGIII, hlyA, cnf1, fimH*	[13]
AIEC21	I	I-CD (colon)	Spain	+	17.0	± 7.8	0.109	± 0.013	1297.1	± 625.2	O6:H1	B2	+	−	*papC, papGII, iucD, hlyA, cnf1, fimH*	[13]
AIEC12	I	IC-CD (colon)	Spain	+	22.3	± 3.9	0.142	± 0.017	93.697	± 55.93	O26:H−	B2	+	−	*papC, afa/draBC, iucD, fimH*	[13]
AIEC20	I	IC-CD (ileum)	Spain	+	14.2	± 6.2	0.125	± 0.098	343.89	± 244.6	O11:H18	D	+	−	*papC, afa/draBC, iucD, fimH*	[13]
AIEC17	I	I-CD (ileum+colon)	Spain	+	21.6	± 17.5	0.266	± 0.055	1053	± 75	ONT:HNT	D	+	−	*neuC, ibeA, fimH*	[13]
AIEC05	I	CD (ileum+colon)	Spain	+	9.4	± 2.2	0.202	± 0.042	704.91	± 714	O1:H−	B2	+	−	*papC, papGII, neuC, fimH*	[13]
AIEC02	I	CD (colon)	Spain	+	0.9	± 1.0	0.802	± 0.035	2187.8	± 4.794	O8:H21	B2	+	−	*papC, neuC, fimH*	[13]
AIEC01	I	I-CD (ileum)	Spain	+	15.9	± 9.3	0.284	± 0.106	1566.7	± 1060	O6:H1	B2	+	−	*sfa/focDE, iucD, fimH*	[13]
AIEC09	I	IC-CD (colon)	Spain	+	5.4	± 4.0	0.216	± 0.010	2562.3	± 240.6	ONT:H−	B2	+	−	*papC, sfa/focDE, iucD, hlyA, cnf1, fimH, fimAv$_{MT78}$*	[13]
AIEC24	I	IC-CD (colon)	Spain	+	2.0	± 1.4	0.309	± 0.138	1625.6	± 115.6	ONT:H−	A	+	−	*iucD, fimH, fimAv$_{MT78}$*	[13]
AIEC23	I	C-CD (ileum)	Spain	+	9.7	± 0.7	0.568	± 0.148	2362.1	± 250.2	O5:HNT	A	−	−	*fimH*	[13]
AIEC11	I	I-CD (ileum)	Spain	+	4.4	± 3.4	0.508	± 0.081	847.95	± 512.8	O22:H1	B2	+	−	*afa/draBC, iucD, fimH*	[13]
AIEC15-1	I	I-CD (ileum)	Spain	+	10.0	± 1.4	0.305	± 0.159	659.75	± 437	O22:H1	B2	−	−	*fimH*	[13]
AIEC14-1	I	I-CD (ileum+colon)	Spain	+	9.8	± 5.2	0.238	± 0.011	800.69	± 252.4	O22:H1	B2	−	−	*fimH*	[13]
AIEC16-2	I	I-CD (ileum)	Spain	+	9.7	± 3.6	1.400	± 0.424	921.05	± 489.7	O22:H1	B2	−	−	*fimH*	[13]
AIEC13	I	UC (colon)	Spain	+	7.9	± 4.3	1.400	± 0.000	225.91	± 181.6	O25:H4	B2	+	−	*papC, papGIII, iucD, hlyA, cnf1, ibeA, fimH*	[13]
AIEC19	I	non-IBD (colon)	Spain	+	2.4	± 0.7	0.111	± 0.016	1568.1	± 1726	ONT:H−	A	+	−	*iucD, fimH, fimAv$_{MT78}$*	[13]
AIEC07	I	non-IBD (ileum)	Spain	+	20.0	± 13.4	0.565	± 0.392	1692.6	± 296.8	O22:H7	B1	+	−	*papC, iucD, fimH*	[13]
AIEC04	I	non-IBD (ileum)	Spain	+	21.6	± 8.9	0.320	± 0.016	584.69	± 418.5	O6:HNT	B2	+	−	*papC, sfa/focDE, iucD, hlyA, cnf1, fimH, fimAv$_{MT78}$*	[13]
AIEC10	I	non-IBD (ileum)	Spain	+	5.9	± 1.0	0.226	± 0.192	1413.7	± 51.37	O159:H34	A	−	−	*fimH*	[13]
AIEC06	I	non-IBD (colon)	Spain	+	10.2	± 3.4	0.177	± 0.019	1717.7	± 307.9	O6:H5	B2	+	−	*papC, papGIII, sfa/focDE, hlyA, cnf1, fimH*	[13]
AIEC08	I	non-IBD (colon)	Spain	+	1.1	± 0.2	0.172	± 0.066	104.75	± 49.71	O25:H4	B2	+	−	*papC, papGIII, iucD, ibeA, fimH*	[13]
LF82	I	I-CD (ileum)	France	+	25.7	± 15.7	2.261	± 1.349	776.88	± 304.8	O83:H1	B2	+	−	*ibeA, fimH, fimAv$_{MT78}$*	[12]

Statistical analyses

The Fisher's exact test (small contingency tables) or Pearson's X^2 test (frequencies higher than five within cells) was used to measure the significance of frequency values using SPSS 15.0 software.

Correspondence analysis was used to determine if a particular distribution of virulence-associated genes correlated with the serogroup, phylogroup, AIEC phenotype, ExPEC-like genotype (more than two virulence genes in addition to *fimH*), origin of the strains (extraintestinal/intestinal), and/or disease caused (intra-abdominal pus, wound infection, sepsis, meningitis, urinary tract infections (UTIs), and IBD). The input variables were the presence/absence of virulence genes: *papC, sfa/focDE, afa/draBC, hlyA*, *cnf1, iucD, neuC, ibeA, fimH*, and $fimAv_{MT78}$) and all 86 *E. coli* strains were included in the analysis. Correspondence analysis was performed with the CANOCO programme (version 4.5 for Windows) using biplot scaling [29]. To corroborate the significance of the dispersion of the samples in the plot according to their serogroup, phylogroup, AIEC phenotype, ExPEC-like genotype and origin of the strains, an ANOVA test was applied using Tukey's Post-Hoc test for multi-comparisons of those variables comprising more than two subgroups of samples. For quantitative variables, such as adhesion (I_ADH), invasion (I_INV), and intra-macrophage replication (I_REPL) indices, the Pearson correlation coefficient was used.

Results

Presence of AIEC-like strains amongst ExPEC strains

The genetic and phenotypic characteristics of the 63 ExPEC and 23 intestinal AIEC strains used in this study are listed in Table 20 Strains belonging to serogroups O1 (n=3), O2 (n=2), O6 (n=21), O7 (n=1), O16 (n=1), O18 (n=5), O25 (n=16), O45 (n=3), O51 (n=1), O75 (n=2), O83 (n=6), O101 (n=1), and ONT (n=1), which were obtained from extraintestinal infections, were selected to be compared with a collection of intestinal AIEC strains belonging to serogroups O1 (n=1), O5 (n=1), O6 (n=5), O8 (n=1), O11 (n=1), O22 (n=5), O25 (n=2), O26 (n=1), O83 (n=1), O159 (n=1), and ONT (n=4).

After determining the capacity of ExPEC strains to adhere to and invade intestinal epithelial cells and their ability to survive and replicate within macrophages, we classified four strains (6.35%) as AIEC strains (Table 21). These strains are referred to as "extraintestinal AIEC" in this study. Two of these strains were

isolated from patients suffering from sepsis, and the other two stains were isolated from UTIs. The extraintestinal AIEC strains belonged to the O6:H1 (two strains), O25:H4, and O83:H1 serotypes. These serotypes comprised 21.7%, 8.7%, and 4.3% of intestinal AIEC strains, respectively.

Thus, the majority of the ExPEC strains that were tested did not exhibit the phenotypic features that characterise the AIEC pathovar.

Table 21. Frequency of ExPEC strains with AIEC phenotype.

	Total	AIEC frequency	
Extraintestinal infection	N	N	(%)
Intra-abdominal pus	1	0	0
Meningitis	12	0	0
Sepsis	21	2	9.5
Urinary tract infections	28	2	7.1
Wound infection	1	0	0
	63	4	6.35

Distribution of virulence genes in AIEC and ExPEC strains

The distribution of virulence-associated genes in ExPEC strains was similar to that obtained for AIEC strains isolated from human intestinal mucosa with the exception of the *sfa/focDE* operons, which were more prevalent amongst ExPEC strains ($P = 0.013$) (Table 22). The distribution of phylogroups was also similar, with B2 being the most abundant phylogroup (85.7% and 69.6% of ExPEC and AIEC strains, respectively). Regarding the AIEC strains, all of the strains studied in the present report harboured the *fimH* gene. The *papC* and *iucD* genes were also prevalent, being present in more than 50% of the AIEC strains. The *papGII* and *papGIII* alleles were the most frequent alleles found amongst ExPEC and AIEC strains.

Correspondence analysis for the presence of virulence genes in the strains corroborated these observations (Figure 28). Neither the intestinal/extraintestinal origin nor the AIEC phenotype was able to explain the segregation of the strains, thus indicating that AIEC and ExPEC pathovars had similar genotypes (Figure 28A). Moreover, no correlation with adhesion, invasion, or intra-macrophage replication indices was detected. Similarly, no segregation was observed between strains that caused different diseases or between strains with distinct phylogenetic origins. A more representative collection of strains from all phylogroups and from all types of extraintestinal infections would be necessary to corroborate this observation. The virulence gene profiles of the strains were primarily associated with the serogroup, as shown in Figure 28B. The majority of O6 strains appeared segregated from the O83 and O25 serogroups by Axis 1 ($P < 0.001$), whereas

Axis 2 separated the O83 strains from the majority of O6 and O25 strains ($P < 0.001$). These results indicate that O6 and O83 strains clearly clustered separately in the CA analysis by their virulence gene profile, whereas O25 strains showed a higher variability of virulence gene sets. Two main clusters of O25 strains appeared in the CA plot. The one situated in the upper-right side of the plot grouped *afa/draBC* positive O25 strains, whereas the O25 strains clustering closer to O6 and O83 strains were *afa/draBC* negative. In particular, those virulence genes that had a better correlation with the O6 serogroup were *hlyA* with a prevalence of 80.8% within the serogroup, *cnf1* and *sfa/focDE*, each with a prevalence of 76.9%, and *papC*, with a prevalence of 65.4%. Amongst the O83 strains, 100% were positive for *ibeA* and 71.4% were positive for $fimAv_{MT78}$. Finally, *iucD* and *ibeA* were present in 83.3% and 50% of the O25 strains, respectively. Moreover, six out of the nine (66.7%) strains that were positive for *afa/draBC* belonged to the O25 serogroup ($P = 0.003$).

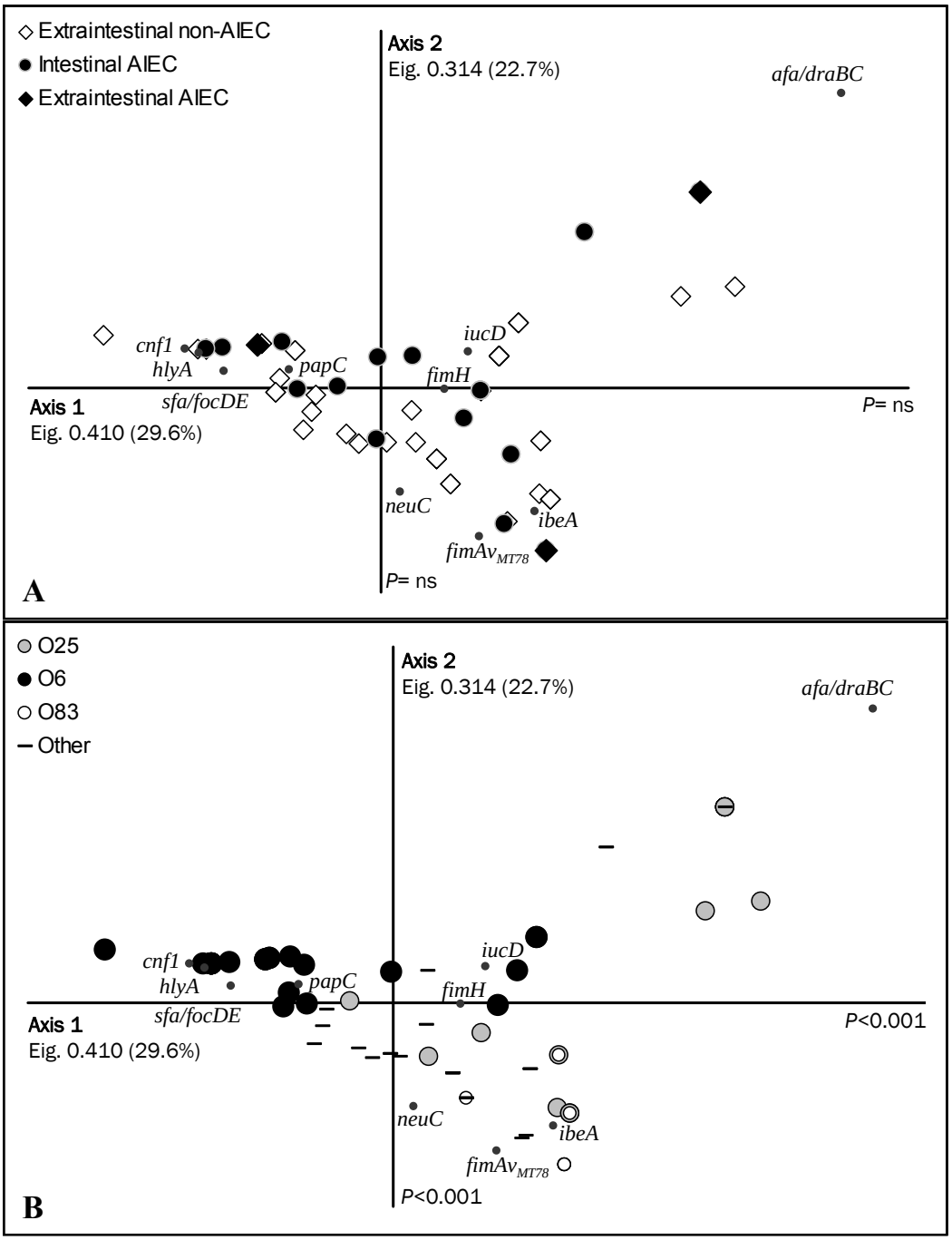
Table 22. Frequency of virulence-associated genes by phenotype (AIEC, non-AIEC) and origin (extraintestinal, intestinal).

Virulence gene	non-AIEC (n=59)	AIEC (n=27)	*P*	Extraintestinal (n=63)	Intestinal (n=23)	*P*
papC	25 (42.4)	14 (51.9)	NS	27 (42.9)	12 (52.2)	NS
papGI	1 (4.0)†	0	NS	1 (3.7)	0	NS
papGII	6 (24)	3 (21.4)	NS	7 (25.9)	2 (16.7)	NS
papGIII	11 (44)	4 (28.6)	NS	11 (40.7)	4 (33.3)	NS
papGI-II	2 (8.0)	0	NS	2 (7.4)	0	NS
papGII-III	2 (8.0)	0	NS	2 (7.4)	0	NS
sfa/focDE	27 (45.8)	6 (22.2)	0.031	29 (46.0)	4 (17.4)	0.013
afa/draBC	5 (8.5)	4 (14.8)	NS	6 (9.5)	3 (13.0)	NS
fimH	54 (91.5)	27 (100.0)	NS	58 (92.1)	23 (100.0)	NS
$fimAv_{MT78}$	16 (27.1)	6 (22.2)	NS	17 (27.0)	5 (21.7)	NS
neuC	15 (25.4)	3 (11.1)	NS	15 (23.8)	3 (13.0)	NS
iucD	39 (66.1)	15 (55.6)	NS	42 (66.7)	12 (52.2)	NS
ibeA	19 (32.2)	5 (18.5)	NS	20 (31.7)	4 (17.4)	NS
hlyA	25 (42.4)	8 (29.6)	NS	27 (42.9)	6 (26.1)	NS
cnf1	21 (35.6)	8 (29.6)	NS	23 (36.5)	6 (26.1)	NS

'AIEC' includes intestinal and extraintestinal AIEC; 'non-AIEC' includes only ExPEC strains; 'Extraintestinal' includes AIEC and non-AIEC ExPEC strains; 'Intestinal' includes only intestinal AIEC strains.
†For papG alleles, percentages are calculated with respect to papC positive strains.
Percentages are indicated in parentheses.

Although they showed distinct phenotypes (in terms of adhesion, invasion, and intracellular replication abilities), the AIEC and ExPEC strains shared similar serotypes, phylogenetic origins, and virulence-associated gene distributions.

Figure 28. Correspondence analysis of the distribution of 10 virulence-associated genes (*papC*, *sfa/focDE*, *afa/draBC*, *fimH*, *fimAv*$_{MT78}$, *neuC*, *iucD*, *ibeA*, *cnf1*, and *hlyA*) in 63 ExPEC strains and 23 intestinal AIEC strains.



Eigenvalues (Eig.) and percentages of variance are provided for each axis. **A**: Extraintestinal/intestinal origin of the strains and AIEC phenotype are specified in this plot. **B**: Serogroup was the sole factor that explained the segregation of the strains (only the most frequent serogroups in our collection [O6, O25 and O83] are specified). Axis 1 explains the segregation of O6 strains from the strains belonging to the O83 and O25 serogroups (P<0.001), whereas axis 2 segregated O83 strains from the O6 and O25 serogroups (P<0.001).

Clonality and phylogenetic relationships amongst O6:H1, O25:H4, and O83:H1 extraintestinal and intestinal AIEC strains.

Multilocus sequence typing (MLST) is a DNA sequencing-based method that has become a popular tool for characterising pathogenic microorganisms including *E. coli* [28]. Using MLST, the genetic relatedness of isolates can be compared, and closely related organisms can be grouped together in clonal complexes. We compared the four extraintestinal AIEC strains with five intestinal AIEC strains of identical serotypes using MLST in order to check for a possible phylogenetic relationship amongst them. Interestingly, the strains segregated into three distinct sequence types (ST) according to their serotype, irrespective of their intestinal/extraintestinal origin. In particular, the CD-associated strains AIEC01 and AIEC21, the UPEC OL96a strain, and the PP215 sepsis-associated strain all belonged to the O6:H1 serotype and the B2 phylogroup, and they carried the same combination of alleles across the seven sequenced loci corresponding to ST73 of the ST73 Clpx. Additionally, AIEC LF82, isolated from a CD patient, and the septicemic strain PP16 belonged to phylogroup B2, ST135 (no Clpx association). Finally, two intestinal O25:H4 AIEC strains isolated from a UC patient (AIEC13) and a non-IBD control (AIEC08) and the UPEC FV7563 strain (O25:H4 CTMX-15 positive) all belonged to phylogroup B2 and displayed ST131 (no Clpx association).

As shown in Figure 29, all of the intestinal and extraintestinal AIEC strains belonging to the O6:H1 (ST73), O83:H1 (ST135), and O25:H4 (ST131) serotypes (and ST types) harboured the pathogenicity-associated island marker *malX* and the uropathogenic-specific protein *usp*, and they all possessed a group II polysaccharide capsule (*kpsM-II*). In contrast, the secreted autotransporter toxin (*sat*) gene was detected in the four AIEC strains with O6:H1 serotype (ST73) and also in one O25:H4 (ST131) extraintestinal AIEC strain. The serum-resistance associated gene (*traT*) was identified in three AIEC strains belonging to O6:H1 serotype (ST73) and in two intestinal AIEC strains with O25:H4 serotype (ST131).

We compared the *XbaI* PFGE macrorestriction profiles of the intestinal and extraintestinal AIEC strains sharing the same ST and phylogroup. PFGE is a highly discriminatory method and is useful for detecting small DNA differences that rapidly accumulate in the bacterial genome. We used this tool to better differentiate the compared strains by identifying clusters with different similarity values. As expected, most strains of the same serotype, phylogenetic group, and ST grouped together in the dendrogram (Figure 29). Thus, the macrorestriction analysis demonstrated that the four strains of serotype O6:H1 B2 ST73 clustered

together with 69.8% similarity. In particular, OL96a, AIEC21, and AIEC01 grouped with 74.6% similarity. The two O83:H1 B2 ST135 strains (intestinal LF82 and ExPEC PP16) exhibited a similarity value of 77.8%. Finally, the two intestinal O25:H4 B2 ST131 strains (AIEC08 and AIEC13) grouped together in the dendrogram (75% similarity), while the UPEC FV7563 isolate (CT-X-M15 *afa/draBC*) appeared to be very different, exhibiting only 48.1% similarity.

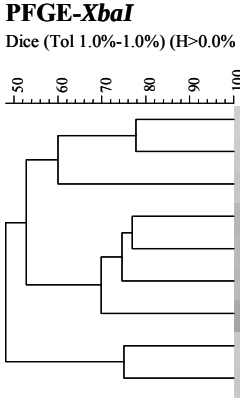
Further PFGE analysis introducing additional intestinal and extraintestinal O83 strains demonstrated that a diversity of pulsotypes existed amongst this serogroup, which segregated according to their flagellar H type and virulence genotype. Thus, the six strains of serotype O83:H1 (including intestinal and extraintestinal AIEC strains and intestinal and extraintestinal non-AIEC strains) grouped together with 75.2% similarity (Figure 30). Two clusters with similarities >85% displayed a close genetic relationship; in particular, the AIEC strain LF82 clustered together with the sepsis-associated strain H102A with 86.5% similarity, and the ECG043 intestinal strain clustered with the UTI strain OB79A (88.2% similarity).

Therefore, although the majority of ExPEC strains did not exhibit an AIEC phenotype, a minority of strains that did have this phenotype were genetically related to some intestinal AIEC strains, as revealed by MLST and, for certain strains, by PFGE.

Discussion

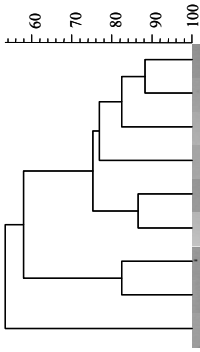
Despite the AIEC pathovar has been repeatedly associated with CD [1, 5, 12-14], some uncertainty exists regarding: i) the genetic relationship between AIEC strains and other pathogenic and non-pathogenic *E. coli* strains; ii) its particular identity as pathovar, and iii) the putative involvement of AIEC strains in extraintestinal diseases in addition to their suspected role in inflammatory bowel disease. For that reason, the aim of this work was to determine the AIEC phenotype of a collection of ExPEC strains and further search for a common phylogenetic origin for the intestinal and extraintestinal AIEC strains.

Figure 29. Consensus UPGMA dendrogram generated from the Dice coefficients of *XbaI* PFGE profiles of the four extraintestinal AIEC detected in this study (OL96a, PP215, PP16, and FV7563) and of the five intestinal AIEC strains of similar serotypes. Serotype, phylogroup, ST type and virulence-associated genes are specified.



Id Strain	Origin	Serotype	Phylo	ST types	Virulence gene profile
LF 82	Intestinal, CD	O83:H1	B2	ST135	*ibeA fimAv$_{MT78}$ fimH malX usp kpsM-II*
PP 16	Sepsis	O83:H1	B2	ST135	*ibeA fimAv$_{MT78}$ fimH malX usp kpsM-II*
FV7563	UTI	O25:H4	B2	ST131	*afa/draBC iucD fimH sat malX usp kpsM-II*
OL 96a	UTI	O6:H1	B2	ST73	*papGII sfa/focDE(focG) iucD hlyA cnf1 fimH sat malX usp iroN kpsM-II*
AIEC 01	Intestinal, CD	O6:H1	B2	ST73	*sfa/focDE(focG) iucD fimH sat traT malX usp iroN kpsM-II*
AIEC 21	Intestinal, CD	O6:H1	B2	ST73	*papGII iucD hlyA cnf1 fimH sat traT malX usp kpsM-II*
PP 215	Sepsis	O6:H1	B2	ST73	*papGII sfa/focDE(focG) iucD hlyA cnf1 fimH sat traT malX usp iroN kpsM-II*
AIEC 08	Intest., non-IBD	O25:H4	B2	ST131	*papGIII ibeA iucD fimH traT malX usp cdtB kpsM-II*
AIEC 13	Intestinal, UC	O25:H4	B2	ST131	*papGIII ibeA iucD hlyA cnf1 fimH cvaC traT malX usp iroN cdtB kpsM-II*

Figure 30. Consensus UPGMA dendrogram generated from the Dice coefficients of *XbaI* PFGE profiles of six O83 ExPEC strains and three O83 intestinal *E. coli* strains.



Id Strain	Origin	AIEC	Serotype	Phylo	Virulence gene profile
ECG043	Intestinal, non-IBD	–	O83:H1	B2	*ibeA fimH fimAv$_{MT78}$ malX usp kpsM-II*
OB79A	UTI	–	O83:H1	B2	*ibeA fimH fimAv$_{MT78}$ malX usp kpsM-II*
PP16	Sepsis	+	O83:H1	B2	*ibeA fimH fimAv$_{MT78}$ malX usp kpsM-II*
OB59A	UTI	–	O83:H1	B2	*ibeA fimH malX usp kpsM-II*
H102A	Sepsis	–	O83:H1	B2	*ibeA fimH fimAv$_{MT78}$ malX usp kpsM-II*
LF82	Intestinal, CD	+	O83:H1	B2	*ibeA fimH fimAv$_{MT78}$ malX usp kpsM-II*
ECG009	Intestinal, CD	–	O83:H–	B2	*sfa/focDE neuC iucD ibeA fimH fimAv$_{MT78}$ cdtB malX usp iroN kpsM-II*
SM57	Meningitis	–	O83:H7	B2	*sfa/focDE neuC iucD ibeA fimH fimAv$_{MT78}$ cdtB tsh cvaC traT malX usp ir*
H126	Sepsis	–	O83:H31	B2	*iucD ibeA fimH cvaC malX usp iroN kpsM-II*

Abbreviations of Figures 29 and 30: CD: Crohn's disease; UC: ulcerative colitis; non-IBD: controls without inflammatory bowel disease; UTI: urinary tract infection. The ECG-043 and ECG-009 strains were used only in this section; their characteristics are described elsewhere [13].

Given the genetic similarity between the AIEC and ExPEC strains with regard to their virulence gene profiles and phylogenetic origins (mainly B2 and D phylogroups [30, 31]), we suspected that a high proportion of ExPEC strains could also be classified as AIEC strains, but only 4 out of 63 (6.35%) ExPEC strains from our collection were found to share the phenotypic characteristics that describe the AIEC pathovar. These results suggest that the AIEC pathovar comprise a particular group of *E. coli* that is closely related to the ExPEC pathovar but is distinguishable by phenotypic traits [1], which give to the pathovar a particular identity. Unfortunately, no specific genes involved in the adhesion, invasion or intra-macrophage replication abilities of AIEC strains have been discovered to date. Although some genes and regulatory processes have been implicated in the pathogenesis of the prototypic AIEC strain LF82 [11, 32-38], most of these genes are present in the non-pathogenic *E. coli* strain K12, thus indicating that differences in gene expression or small sequence variations of these genes might contribute to the AIEC phenotype.

A high diversity of serotypes and virulence gene profiles exists amongst ExPEC strains, which complicates their classification into pathotypes. Although correspondence analysis segregated the strains by their serogroup, AIEC and non-AIEC strains of intestinal and extraintestinal origin were present in all clusters, thus indicating that a variety of seropathotypes can also be found amongst AIEC strains. In particular, those virulence genes that best correlated with O6 strains were *papC*, *sfa/focDE*, *cnf1*, and *hlyA*, whereas *fimAv*$_{TM78}$ and *ibeA* correlated with O83 strains, and *afa/draBC*, *iucD*, and *ibeA* correlated with O25 strains. Nevertheless, some genes were constantly found in all O6, O25 and O83 AIEC strains, both of intestinal and extraintestinal origin (*malX*, *usp*, and *kpsM-II*). These genes have been already described for AIEC strain LF82 [39].

Several studies providing a complete description of the virulence-associated genes of a variety of AIEC strains have been published to date, and all coincide that the AIEC pathovar shows homology to human ExPEC strains [1, 5, 12, 13, 34, 39, 40]. The virulence genes *fimH*, *fimAv*$_{MT78}$, *lpfA*, *papC*, *papGII*, *afaB-afaC*, *sfa/focDE*, ColV plasmid, *iucD*, *iss*, *kpsMII*, *neuC*, *ibeA*, *malX*, *usp*, *chuA*, *hlyA*, *cnf1*, and UPEC PAIs IV$_{536}$,VI$_{536}$, I$_{CFT073}$, II$_{CFT073}$, characteristic of ExPEC strains, have been detected at distinct frequencies in AIEC strains. In addition, virulence genes of other pathogenic *Enterobacteriaceae*, such as *Salmonella* (*ratA*), *Yersinia* (pMT1, *fyuA*, *irp*1 and 2) and *Vibrio* (*hcp*), have been detected in LF82 and other AIEC strains [1]. The presence and prevalence of *papC*, *afa/draBC*, and *fimH* within the AIEC collection used in this study agrees with

previous descriptions of AIEC or intra-mucosal *E. coli* strains isolated by other researchers [1, 5, 12]. In contrast, whereas some AIEC strains in this study carried *sfa/focDE*, *cnf1*, and *hlyA*, these virulence genes have not been detected in other collections of invasive *E. coli* strains [5]. In particular, the virulence factors *hlyD/cnf1* (PAI II_{J96}) have been reported to be present in the genome of 40% of *E. coli* strains isolated from colorectal cancer, whereas it was absent in 8 strains isolated from CD patients [40]. In contrast, we detected six *cnf1* and *hlyA* positive AIEC strains, five of which were isolated from colon specimens and one from the ileum of a healthy individual. The heterogeneity of gene profiles found in different studies can be explained by the great genetic diversity amongst AIEC strains, by the fact that patients came from different geographical regions, or because the *E. coli* collections used are not representative enough of the real *E. coli* diversity present in CD patients. Nevertheless, such genes have been also detected in non-pathogenic *E. coli* and are supposed to actually be contributing to fitness or colonization [41].

A portion of AIEC strains, including the prototype AIEC LF82, showed virulence genes (*fimAv*$_{MT78}$, *neuC*, *ibeA* or *cdt*) that are frequent amongst avian pathogenic *E. coli* (APEC) strains belonging to the subcluster B2-1 defined by Moulin-Schouleur *et al* [42]. Interestingly, these B2-1 APEC strains were reported to be genetically and phenotypically close to certain human ExPEC as revealed by MLST, serotyping and genotyping. The authors suggested that little or no host specificity exist amongst these groups of human and avian *E. coli* strains, and thus APEC might constitute a zoonotic risk. Because previous reports have already addressed the similarity between both pathovars [1, 34, 40], in conjunction with the fact that AIEC-like strains have been detected in granulomatous Boxer dogs and cow mastitis, determining the distribution of AIEC strains in different healthy and diseased animals is a research area that could contribute in the understanding of AIEC epidemiology, host-specificity and possible routes of transmission.

Noticeably, some strains belonging to the same phylogenetic group, having identical serotypes and virulence gene profiles – for example, the five O83:H1 B2 ST135 strains harbouring *fimH*, *fimAv*$_{MT78}$, *ibeA*, *malX*, *usp*, and *kpsM-II* genes – and having a close genetic relationship as determined PFGE (Figure 30) displayed different adhesion, invasion, and intra-macrophage replication abilities and thus different AIEC phenotypes. Similarly, in a previous study we observed that isolates from a given subject had identical PFGE profiles but differed from their AIEC phenotype [13]. This observation led us to postulate that the AIEC phenotype is achieved by differences in gene expression, the existence of single

nucleotide polymorphisms or the loss or gain of DNA by horizontal gene transfer. We agree with the hypothesis that AIEC strains are members of the ExPEC pathovar, which usually reside the intestine [43], but that have evolved taken advantage of the particular "IBD microenvironment" [44]. However, we would remark that the genetic determinants implicated in the AIEC phenotype are at least not detectable by PFGE, MLST or virulence genotyping of known virulence factors. Baumgart *et al*. [1] and Bronowski *et al*. [40] have performed genome subtraction in order to search for unknown AIEC specific genes. However, these studies were designed to compare strains that are in fact very different from each other (they used as "driver" non-pathogenic *E. coli* and UPEC strains), thus obtaining a large number of subtracted genes in addition to those related with the AIEC phenotype. Given the high genetic variability among *E. coli*, a more targeted discrimination, searching for differences between genetically close strains that only differ on their AIEC phenotype, would probably reduce the number of differences and only those genes most involved in producing the AIEC phenotype would appear in the subtraction library.

It should be emphasised that the four extraintestinal AIEC strains detected in our collection, which belonged to the O6:H1, O25:H4, and O83:H1 serotypes, were found to belong to the same clonal groups as some intestinal AIEC strains with the same serotypes, as revealed by MLST. These results suggest that some intestinal AIEC could cause extraintestinal infections or vice versa. Interestingly, one of the most representative clones from our AIEC collection, O6:H1-ST73, is a frequent cause of urinary tract infections and septicaemia. The possible implication of intestinal pathogenic *E. coli* in extraintestinal infections has been suggested [45]. A recent study reports that 6.9% of the strains from a collection of 225 ExPEC strains exhibited a diffuse-adhering phenotype, which is characteristic of the intestinal pathogenic pathovar DAEC (diffuse-adhering *E. coli*) [46]. Moreover, the authors also detected several virulence genes, principally from EAEC (Enteroaggregative *E. coli*), in some ExPEC strains, thus indicating that certain ExPEC strains may carry virulence properties of diarrhoeagenic *E. coli* (DEC). All of these observations suggest that certain AIEC strains might be involved in extraintestinal infections.

ExPEC strains may cause a wide range of extraintestinal infections, especially in immunocompromised patients and in persons exposed to high infective doses or persons who have co-infections [31]. Whether ExPEC strains should be considered true pathogens or merely opportunistic commensal bacteria remains controversial [31, 47, 48]. The same question is also posed for AIEC strains. The presence of AIEC strains in control subjects suggests that additional host and/or

environmental factors are needed for these bacteria to cause an infection even though these strains should have the virulence features needed to cause disease. Amongst the host factors, Crohn's disease-specific genetic susceptibility loci, such as *NOD2*/*CARD15* (implicated in peptidoglycan recognition, tolerance to bacteria, and bacterial clearance) and the autophagy-related genes *ATG16L1* and *IRGM,* could be involved in the management of AIEC infections [16].

We identified two intestinal AIEC strains and one extraintestinal AIEC strain, all of which were of the O25:H4 serotype, which belonged to the emerging and virulent clonal group ST131 [23, 45, 49-51]. Currently, this clonal pathotype is the most prevalent amongst human ExPEC strains. Nicolas-Chanoine *et al.* [23] recently reported the intercontinental emergence of an ExPEC O25:H4-ST131 clone that produces the extended-spectrum β-lactamase CTX-M-15. In the present study, we found that only one (FV7563) of the seven assayed ExPEC O25:H4-ST131 strains producing CTX-M-15 had an AIEC phenotype. This extraintestinal AIEC strain showed a different virulence profile and a very different macrorestriction profile compared to the two intestinal AIEC strains (AIEC08 and AIEC13) of the same clonal group. This is the first time that strains belonging to clone ST131 have been shown to harbour the *papC*, *papGIII*, *ibeA*, *cnf1*, and *hlyA* genes [23, 49, 51].

To conclude, the present study demonstrated that the ExPEC and AIEC pathovars share similar virulence gene sets and that certain strains are phylogenetically related. However, the majority of ExPEC strains did not behave like AIEC strains, thus confirming that the AIEC pathovar is related to the ExPEC pathovar but possesses virulence-specific features. These observations suggest that the AIEC phenotype might be encoded by unknown virulence factors or might be the result of differential expression of key genes. Further investigation of the genes implicated in the AIEC phenotype is necessary.

References

1. Baumgart, M., et al., Culture independent analysis of ileal mucosa reveals a selective increase in invasive *Escherichia coli* of novel phylogeny relative to depletion of Clostridiales in Crohn's disease involving the ileum. *ISME J* 2007;**1**:403-418.

2. Darfeuille-Michaud, A., et al., Presence of adherent *Escherichia coli* strains in ileal mucosa of patients with Crohn's disease. *Gastroenterology* 1998;**115**:1405-13.

3. Fujita, H., et al., Quantitative analysis of bacterial DNA from *Mycobacteria* spp., *Bacteroides vulgatus*, and *Escherichia coli* in tissue samples from patients with inflammatory bowel diseases. *J Gastroenterol* 2002;**37**:509-516.

4. Keighley, M.R., et al., Influence of inflammatory bowel disease on intestinal microflora. *Gut* 1978;**19**:1099-1104.

5. Martin, H.M., et al., Enhanced *Escherichia coli* adherence and invasion in Crohn's disease and colon cancer. *Gastroenterology* 2004;**127**:80-93.

6. Neut, C., et al., Impaired bacterial flora in human excluded colon. *Gut* 1989;**30**:1094-1098.

7. Swidsinski, A., et al., Active Crohn's disease and ulcerative colitis can be specifically diagnosed and monitored based on the biostructure of the fecal flora. *Inflamm. Bowel Dis.* 2008;**14**:147-161.

8. Swidsinski, A., et al., Spatial organization and composition of the mucosal flora in patients with inflammatory bowel disease. *J. Clin. Microbiol.* 2005;**43**:3380-3389.

9. Willing, B., et al., Twin studies reveal specific imbalances in the mucosa-associated microbiota of patients with ileal Crohn's disease. *Inflamm. Bowel Dis.* 2009;**15**:653-60.

10. Kaper, J.B., J.P. Nataro, and H.L.T. Mobley, Pathogenic *Escherichia coli. Nat Rev Micro* 2004;**2**:123-140.

11. Boudeau, J., et al., Invasive ability of an *Escherichia coli* strain isolated from the ileal mucosa of a patient with Crohn's disease. *Infect. Immun.* 1999;**67**:4499-4509.

12. Darfeuille-Michaud, A., et al., High prevalence of adherent-invasive *Escherichia coli* associated with ileal mucosa in Crohn's disease. *Gastroenterology* 2004;**127**:412-421.

13. Martinez-Medina, M., et al., Molecular diversity of *Escherichia coli* in the human gut: new ecological evidence supporting the role of adherent-invasive *E. coli* (AIEC) in Crohn's disease. *Inflamm Bowel Dis* 2009;**15**:872-882.

14. Sasaki, M., et al., Invasive *Escherichia coli* are a feature of Crohn's disease. *Lab Invest* 2007;**87**:1042-1054.

15. Glasser, A.-L., et al., Adherent invasive *Escherichia coli* strains from patients with Crohn's disease survive and replicate within macrophages without inducing host cell death. *Infect. Immun.* 2001;**69**:5529-5537.

16. Glasser, A.-L. and A. Darfeuille-Michaud, Abnormalities in the handling of intracellular bacteria in Crohn's disease: a link between infectious etiology and host genetic susceptibility. *Arch. Immunol. Ther. Exp.* 2008;**56**:237-244.

17. Masseret, E., et al., Genetically related *Escherichia coli* strains associated with Crohn's disease. *Gut* 2001;**48**:320-325.

18. Bidet, P., et al., Combined Multilocus Sequence Typing and O serogrouping distinguishes *Escherichia coli* subtypes associated with infant urosepsis and/or meningitis. *J. Infect. Dis.* 2007;**196**:297-303.

19. Blanco, M., et al., Virulence factors and O groups of *Escherichia coli* strains isolated from cultures of blood specimens from urosepsis and non-urosepsis patients. *Microbiologia* 1994 **10**:249-56.

20. Blanco, M., et al., Virulence factors and O groups of *Escherichia coli* isolates from patients with acute pyelonephritis, cystitis and asymptomatic bacteriuria. *Eur J Epidemiol.* 1996 **12**:191-8.

21. Moulin-Schouleur, M., et al., Common virulence factors and genetic relationships between O18:K1:H7 *Escherichia coli* isolates of human and avian origin. *J. Clin. Microbiol.* 2006;**44**:3484-3492.

22. Naves, P., et al., Correlation between virulence factors and in vitro biofilm formation by *Escherichia coli* strains. *Microb. Pathog.* 2008;**45**:86-91.

23. Nicolas-Chanoine, M.-H., et al., Intercontinental emergence of *Escherichia coli* clone O25:H4-ST131 producing CTX-M-15. *J. Antimicrob. Chemother.* 2008;**61**:273-281.

24. Clermont, O., S. Bonacorsi, and E. Bingen, Rapid and simple determination of the *Escherichia coli* phylogenetic group. *Appl. Environ. Microbiol.* 2000;**66**:4555-4558.

25. Mora, A., et al., Extraintestinal pathogenic *Escherichia coli* O1:K1:H7/NM from human and avian origin: detection of clonal groups B2 ST95 and D ST59 with different host distribution. *BMC Microbiology* 2009;**9**:132.

26. Guinée, P.A., C.M. Agterberg, and W.H. Jansen, *Escherichia coli* O antigen typing by means of a mechanized microtechnique. *Appl Microbiol* 1972;**24**:127-31.

27. CDCPulseNetUSA, One-day (24-28 h) standardized laboratory protocol for molecular subtyping of *Escherichia coli* O157:H7, non-typhoidal *Salmonella* serotypes, and *Shigella sonnei* by pulsed field gel electrophoresis (PFGE). *PulseNet PFGE manual* 2004;**5.1-5.3**:1-13.

28. Wirth, T., et al., Sex and virulence in *Escherichia coli*: an evolutionary perspective. *Mol Microbiol* 2006;**60**:1136-1151.

29. ter Braak, C.J.F. and P. Smilauer, CANOCO Reference manual and user's guide to Canoco for Windows: Software for canonical community ordination (Version 4). 1998;

30. Kotlowski, R., et al., High prevalence of *Escherichia coli* belonging to the B2+D phylogenetic group in inflammatory bowel disease. *Gut* 2007;**56**:669-675.

31. Russo, T.A. and J.R. Johnson, Proposal for a new inclusive designation for extraintestinal pathogenic isolates of *Escherichia coli*: ExPEC. *J. Infect. Dis.* 2000;**181**:1753-1754.

32. Barnich, N., et al., Regulatory and functional co-operation of flagella and type 1 pili in adhesive and invasive abilities of AIEC strain LF82 isolated from a patient with Crohn's disease. *Mol. Microbiol.* 2003;**48**:781-794.

33. Barnich, N., et al., Involvement of lipoprotein NlpI in the virulence of adherent invasive *Escherichia coli* strain LF82 isolated from a patient with Crohn's disease. *Infect. Immun.* 2004;**72**:2484-2493.

34. Boudeau, J., N. Barnich, and A. Darfeuille-Michaud, Type 1 pili-mediated adherence of *Escherichia coli* strain LF82 isolated from Crohn's disease is involved in bacterial invasion of intestinal epithelial cells. *Mol. Microbiol.* 2001;**39**:1272-1284.

35. Bringer, M.-A., et al., HtrA stress protein is involved in intramacrophagic replication of adherent and invasive *Escherichia coli* strain LF82 isolated from a patient with Crohn's disease. *Infect. Immun.* 2005;**73**:712-721.

36. Bringer, M.-A., et al., The oxidoreductase DsbA plays a key role in the ability of the Crohn's disease-associated adherent-invasive *Escherichia coli* strain LF82 to resist macrophage killing. *J. Bacteriol.* 2007;**189**:4860-4871.

37. Rolhion, N., et al., Strong decrease in invasive ability and outer membrane vesicle release in Crohn's disease-associated adherent-invasive *Escherichia coli* strain LF82 with the *yfgL* gene deleted. *J. Bacteriol.* 2005;**187**:2286-2296.

38. Rolhion, N., F.A. Carvalho, and A. Darfeuille-Michaud, OmpC and the sigma(E) regulatory pathway are involved in adhesion and invasion of the Crohn's disease-associated *Escherichia coli* strain LF82. *Mol. Microbiol.* 2007;**63**:1684-1700.

39. Simpson, K.W., et al., Adherent and invasive *Escherichia coli* is associated with granulomatous colitis in Boxer dogs. *Infect. Immun.* 2006;**74**:4778-4792.

40. Bronowski, C., et al., A subset of mucosa-associated *Escherichia coli* isolates from patients with colon cancer, but not Crohn's disease, share pathogenicity islands with urinary pathogenic *E. coli*. *Microbiology* 2008;**154**:571-583.

41. Hejnova, J., et al., Characterization of the flexible genome complement of the commensal *Escherichia coli* strain A0 34/86 (O83 : K24 : H31). *Microbiology* 2005;**151**:385-398.

42. Moulin-Schouleur, M., et al., Extraintestinal pathogenic *Escherichia coli* strains of avian and human origin: link between phylogenetic relationships and common virulence patterns. *J. Clin. Microbiol.* 2007;**45**:3366-3376.

43. Russo, T.A., et al., Chromosomal restriction fragment length polymorphism analysis of *Escherichia coli* strains causing recurrent urinary tract infections in young women. *J Infect Dis* 1995 **172**:440-5.

44. Sepehri, S., et al., Phylogenetic analysis of inflammatory bowel disease associated *Escherichia coli* and the *fimH* virulence determinant. *Inflamm Bow Dis* 2009;**DOI 10.1002/ibd.20966**:

45. Manges, A., et al., Endemic and epidemic lineages of *Escherichia coli* that cause urinary tract infections. *Emerg Infect Dis* 2008 **14**:1575-83.

46. Abe, C.M., et al., Uropathogenic *Escherichia coli* (UPEC) strains may carry virulence properties of diarrhoeagenic *E. coli*. *FEMS Immunol. Med. Microbiol.* 2008;**52**:397-406.

47. Dobrindt, U., (Patho-)Genomics of *Escherichia coli*. *Int J Med Microbiol* 2005 **295**:357-71.

48. Mokady, D., U. Gophna, and E. Ron, Virulence factors of septicemic *Escherichia coli* strains. *Int J Med Microbiol* 2005 **295**:455-62.

49. Blanco, M., et al., Molecular epidemiology of *Escherichia coli* producing extended-spectrum {beta}-lactamases in Lugo (Spain): dissemination of clone O25b:H4-ST131 producing CTX-M-15. *J. Antimicrob. Chemother.* 2009;**63**:1135-1141.

50. Cagnacci, S., et al., European emergence of ciprofloxacin-resistant *Escherichia coli* clonal groups O25:H4-ST 131 and O15:K52:H1 causing community-acquired uncomplicated cystitis. *J Clin Microbiol.* 2008 **46**:2605-12.

51. Johnson, J.R., et al., Epidemic clonal groups of *Escherichia coli* as a cause of antimicrobial-resistant urinary tract infections in Canada, 2002 to 2004. *Antimicrob. Agents Chemother.* 2009;**53**:2733-2739.

General discussion

Almost eighty years after Dr. Burrill B. Crohn described Crohn's disease using the term 'Regional Ileitis', its etiology remains uncertain despite the considerable effort has been made since then. Several genetic factors that confer a higher risk to acquire the disease have been described. However, the odds ratios of the susceptibility loci are usually very low, indicating that a low frequency of patients explains these correlations. Many evidences indicate that several environmental factors can also contribute to the disease. However, the complexity of the enteric bacterial community along with the huge variety in the diets and hygiene practices of CD patients, as well as other environmental aspects, make it an extremely complicated task to find a specific pathogen that causes CD. Moreover, not only must the genetic and environmental factors be taken into account, but also the interactions between them. The understanding of host-microbe interactions acquires major relevance in an environment such as the intestine, where large amounts of bacteria are in juxtaposition with cells from the host's epithelium and immune system.

The principal aim of this work was to describe which bacterial populations were present in the intestinal mucosa of CD patients so that, by comparing them with those of control subjects, we could identify putative etiologic agents. Some research has already been done in this direction, but many studies used culture-based techniques [257], fecal samples [100, 255] or paraffin-embedded tissue [187, 216]. Culture is clearly limiting due to the complex nature of the bacterial flora, as well as the fastidious nutritional requirements and strict incubation conditions of a number of species. For this reason, we used a PCR-DGGE-sequencing approach with primers covering the entire *Eubacteria* domain. DGGE is, perhaps, one of the most commonly-used techniques among the culture-independent fingerprinting techniques. It is usually the preferred technique for assessing the structure and dynamics of microbial communities because it allows the processing of a large number of samples in comparison with other molecular techniques such as cloning, which is more precise but also more expensive and time consuming. The microbial composition of feces is known to differ from tissue samples as demonstrated in healthy subjects [90], which indicates that for certain studies, samples of feces are a poor substitute for biopsies. As our aim was to identify putative etiologic agents, we considered it essential to scrutinize the mucosa-associated microbiota because of their proximity to the host cells and immune system. Furthermore, we used fresh biopsies rather than formalin-fixed or paraffin-embedded tissue (from which it is difficult to efficiently obtain pure DNA extracts), thereby reducing the possibility of overlooking minority populations [292].

1 Alterations in the microbial composition of Crohn's disease patients

Microbial fingerprints associated with the intestinal mucosa of CD patients appeared altered with respect to those from non-IBD controls. Moreover, CD patients showed higher patient-to-patient variability of bacterial composition in comparison with control subjects. In turn, the non-IBD controls shared several common bands which, in general, the highest similitude was for *Faecalibacterium prausnitzii* and this constituted a "healthy pattern". Several different clones of *F. prausnitzii* were detected in the intestinal mucosa of this group of subjects. Among CD patients, there was a higher prevalence of *Enterobacteriaceae*, in particular *Escherichia coli*, *Clostridium* spp. (Cluster XIVa) and *Ruminococcus torques*, as well as the presence of uncommon opportunistic pathogens unevenly distributed among patients. As a result of these observations, several lines of investigation were then opened in our laboratory, with the main focus on *F. prausnitzii* and *E. coli*. Those concerning the study of *E. coli* populations are presented in this work and will be discussed later on.

Several independent studies using distinct technical approaches and sample types have demonstrated that the microbial community of CD patients is distinguishable from control subjects, and our approach confirms these observations. Studies using culture-independent methods that have already reported such imbalances include Kleessen *et al.* in 2002 [194] using FISH (*fluorescent in situ hybridisation*), followed by Seksik *et al.* [100] using TTGE (*temporal temperature gradient gel electrophoresis*), Mangin *et al.* [255] (cloning), both using fecal samples, then Ott *et al.* [209] using SSCP (*single strand conformation polymorphism*) and Prindville *et al.* [214] (cloning). Together with and further to the publication of our study [205], many other authors have confirmed the existence of dysbiosis in CD [180, 203, 204, 207, 210-213, 215, 217, 218, 220]. Nevertheless, we still need to pay particular attention to those bacterial species that have been associated with CD patients, in order to elucidate which candidates are of greater relevance to this disease.

1.1 Principal populations responsible for dysbiosis in Crohn's disease

The results obtained by different laboratories concerning the main bacterial populations involved in dysbiosis are discussed in this section. Although the majority of studies report similar results, there are some discrepancies regarding the specific microbiota associated with CD.

1.1.1 *Escherichia coli*

Many works report higher prevalence (qualitative studies) or higher abundance (quantitative studies) of aerobes and facultative anaerobes mostly belonging to γ-*Proteobacteria,* principally *Escherichia coli*, as a feature of a great portion of CD patients. Our results are in agreement with these studies [204, 207, 211, 212, 214, 215, 218-220]. This widely accepted tendency fits well with the discovery of the AIEC pathovar which has been associated with Crohn's disease [288], thus suggesting that **certain strains of *E. coli* might be relevant in CD pathogenesis**. For that reason, we have focused further studies on revising the ecological aspects of *E. coli* populations in CD patients (abundance, richness and diversity) as well as on their pathogenic behaviour.

1.1.2 *Faecalibacterium prausnitzii*

F. prausnitzii belongs to the *Clostridium leptum* subgroup (Cluster IV) of the phylum *Firmicutes* and is one of the most common bacterium in the human gut [293, 294]. Our results coincide with previous observations [210, 214, 216] and have been confirmed by studies elsewhere [204, 207, 211, 215, 218]. Depending on the methodology used, these works have demonstrated imbalances at distinct taxonomic levels, from phylum to species. Nevertheless, the *F. prausnitzii* species has begun to attract much more attention and is the focus of interest of very recent reports [219, 295]. Although this bacterium is present in all subjects, it is not as abundant in CD patients as it is in control patients, particularly for Crohn's ileitis (I-CD) [219]. However, the reduction in *F. prausnitzii* abundance is not CD-specific. Recently, it has been reported that patients suffering from infectious colitis [295] and celiac disease [215] also have reduced levels of this bacterium. In addition, some discrepancies have been found in UC patients. Swidsinski *et al.* [215] found significant differences between these two types of IBD, whereas Sokol *et al.* [295] did not. Probably, such apparently incongruent results are due to distinct methodological approaches and, more specifically, to the different types of samples used (feces and tissue respectively). Interestingly, among CD patients, the levels of *F. prausnitzii* recover slightly during remission [215, 295] and a reduction is associated with a higher risk of postoperative recurrence of ileal CD [296].

Finding the exact cause of the decrease in *F. prausnitzii* and the implications of such an imbalance for CD or other intestinal diseases will require further investigation. However, high-dose cortisol therapy or Infliximab restored the counts of *F. prausnitzii* in CD patients, thus suggesting that this bacterium is

somehow suppressed by the host immune system [215]. Swidsinski and co-workers also pointed out the possibility that this bacterium could be inhibited or displaced by another bacterium such as an opportunistic pathogen.

F. prausnitzii is a gram-negative, non-spore-forming and strictly anaerobic bacterium [297]. Along with other members of *Clostridiaceae* (including *Eubacterium rectale* and *Roseburia*), *F. prausnitzii* forms short-chain fatty acids (SCFAs) during fermentation and these have an important effect on colonic health ([298] and references therein). Of note is butyrate, which serves as an energy source for colonocytes, exerting direct effects upon gene expression in mammalian cells, epithelial barrier integrity and immune regulation; furthermore, it also prevents cancer and ulcerative colitis [298-300]. That is to say, butyrate plays important roles in the metabolism and normal development of colonocytes at several levels, thus contributing to colonic health.

Irrespective of the factors causing the **reduction of *F. prausnitzii*** in CD patients, the implication of this loss to normal intestinal functioning is very important and **may be contributing to the pathogenesis of CD.** Sokol *et al.* [296] have recently reported anti-inflammatory properties of this bacterium using both *in vitro* and *in vivo* models. This bacterium seems to be crucial for gut homeostasis and, for that reason, a study of their applicability as a *probiotic* would be of interest. However, such an objective might be very difficult to achieve since it is highly oxygen sensitive, surviving for less than 2 minutes when exposed to air [297]. It would be easier to design *prebiotics* that would enhance the development of *F. prausnitzii* in the intestine. In fact, the fructan, inulin, has recently been reported to have some stimulatory effect on *F. prausnitzii* [301]; however, this needs further confirmation because contradictory results exist between distinct laboratories [302].

1.1.3 *Clostridium coccoides* (Cluster XIVa)

There are some discrepancies between our study and other reports with respect to *Clostridium coccoides* (subcluster XIVa). We reported a higher prevalence of this taxon in CD patients, which is in agreement with a subsequent study in which *C. coccoides* subgroup was found to be more abundant than *C. leptum* group in CD mucosa [204]. Interestingly, a similar imbalance has been reported for patients suffering from colorectal cancer, which showed higher bacterial diversity index scores for the *C. coccoides* subgroup in comparison with controls [303]. Conversely, however, quite a lot of studies have associated this taxon with control subjects [211, 216-218]. Such inconsistencies could be due to: i)

significant differences in the mean age of patients and controls among studies, since differences in abundance and species diversity have been found in the *C. coccoides* subgroup between adult and elderly subjects [304]; ii) the use of universal or specific primers in molecular-based studies, which can lead to distinct results for a certain sample [304]; iii) the problematic taxonomy of genus *Clostridium*: a marked phylogenetic incoherence for the genus *Clostridium* was reported by Collins *et al*. [305]; iv) the fact that comparisons between CD patients and controls (prevalence, frequency, and/or abundance) are given at a very high taxonomic level.

Due to the complexity of the family *Clostridiaceae*, which contains species that are, not only genetically but also phenotypically, quite distant from each other, any differences found between CD patients and healthy control subjects should be checked at lower levels, with a more precise description given of those species present in both groups of subjects. For example, *Clostridiaceae* contains species which are both gram-positive and gram-negative, sporulating and non-sporulating, mesophiles and thermophiles, and there is a wide range of distinct metabolisms [305]. Moreover, commensal *Clostridiaceae* species include butyrate-producing bacteria which are beneficial to the host and pathogenic bacteria such as *Clostridium colinum*, which is known to cause ulcerative colitis in chickens. In particular, it has been suggested that *C. coccoides* (subcluster XIVa) is a suprageneric cluster [305].

In our study, the species grouped within the *C. coccoides* subgroup were *C. clostridiiforme* (1 OTU), *C. bolteae* (2 OTUs), *C. hathewayi* (1 OTU) – these first three all belonging to the "*C. clostridiiforme* group" [306] –, *C. hylemonae* (2 OTUs), *C. nexile* (1 OTU), and *C. polysaccharolyticum* (1 OTU). In addition, we found 4 OTUs of *C. spiroforme* and 1 OTU of *C. cocleatum*, both belonging to Cluster XVIII [305]. Interestingly, the "*C. clostridiiforme* group" has been identified in several reported cases of bacteremia, rectorectal abscesses and intraperitoneal sepsis ([306] and references therein). These observations lead to think that some species of *Clostridium* could probably be implicated in certain CD complications such as rectorectal abscesses. A very recent study, in which a prospective analysis of *C. difficile* carriage in IBD patients was conducted, reports that this bacterium was more prevalent among IBD patients under clinical remission than non-IBD controls [307]. The authors attributed this higher frequency to the underlying disturbances in innate immunity, to enteric bacterial dysbiosis or to chronic mucosal inflammation and their study supports our results indicating that some species of *Clostridium* may be associated with Crohn's

disease, despite the fact that – unlike *C. coccoides* – *C. difficile* belongs to cluster XI [305].

Considering all the above, the conclusion could be drawn that **certain members of *Clostridium* may act as opportunistic pathogens and may be involved in certain phenotypes of Crohn's disease**. Thus, a detailed study of *Clostridium* diversity at the level of different strains and their pathogenic behaviour in CD might be a central subject to be addressed in further studies.

1.1.4 *Bacteroidetes*

A similar, and in some cases, more pronounced disagreement exists regarding the role of *Bacteroidetes*. Whereas some researchers have found this phylum more prevalent and/or abundant among CD patients compared to non-IBD controls [203, 204, 211, 212], others found no such imbalances or even observed increased abundance in control subjects [209, 213, 217, 218]. In our study, the prevalence of sequences corresponding to *Bacteroidetes* was similar among CD patients and controls. At the species level, we detected two strains of *B. vulgatus* in both CD and non-IBD patients, a unique band identified as *B. thetaiotaomicron* in a control subject and one corresponding to *B. caccae* in two CD patients. With regard to *B. vulgatus*, conflicting results have been published elsewhere. While Conte *et al.* [220] observed lower occurrence of this species in CD and UC patients, Dicksved *et al.* [180] reported a higher abundance in CD patients. The latter also found differences for other species, such as *B. ovatus* and *B. uniformis*, which were associated to CD patients and control subjects respectively. Similar contradictory results have been reported for *B. fragilis* [210, 212-214, 216, 217]. In addition, the *Bacteroides* species in the human gut have been reported to be spatially distributed along an axis from the surface of colonocytes to the luminal contents [308]. Thus, the *Bacteroides* composition can vary depending on the zone sampled, what can be a contributing factor for discrepancies obtained amongst different laboratories.

The work of Dicksved *et al.* [180] is useful because they performed a clone library specific for *Bacteroides* spp. in a study based on twin pairs (concordant and discordant) that suffered from different disease localisations. Interestingly, they observed differences in the composition of *Bacteroides* between CD patients with ileal involvement and those with CD of the colon. The libraries were very different from one subset of patients to another, indicating large interpatient variability. However, it has been shown that a large inter-individual variation exists for *Bacteroides* even in healthy subjects [92].

The phylum *Bacteroidetes* is the second most abundant in the human gastrointestinal tract [86]. *Bacteroides* species, which are saccharolytic gram-negative obligate anaerobes, are commonly considered to play a commensal role. However, some virulent species like *B. fragilis* are frequently isolated from clinical specimens, particularly from abdominal cavity infections. Certain strains carry a variety of mobile genetic elements that might encode virulence factors, or regulate their expression, such as the biosynthesis of capsular polysaccharides and hemolysin proteins [309]. Capsular polysaccharides are known to be involved in the formation of abscesses by *B. fragilis* and might confer an advantageous system for evasion of host immune responses.

Given the above mentioned controversy, **further studies are needed to elucidate whether or not *Bacteroides* species are implicated in CD**, at least in certain phenotypes of CD. These studies should consider: i) separating patients according to disease localisation; ii) working at the species level; iii) studying those species close to the mucosal surface; iv) searching for possible pathogenic features of the strains present in CD mucosa and v) investigating the expression of these pathogenic features by *in vitro* and/or *in vivo* models.

1.1.5 *Ruminococcus torques*

The association of *Ruminococcus torques* with CD patients found in this work cannot be contrasted at present, because of the lack of studies reporting their abundance or prevalence in CD patients. This might be because no significant differences between patients and controls were found or, more probably, despite being a member of the intestinal microbiota [92], it may have passed undetected in previous studies due to the methodological approaches used. However, Prindville *et al.* [214] observed an increased population of *R. gnavus* in the small bowel of CD patients in comparison with controls, which would support our data. Nevertheless, some attention should perhaps be paid to this bacterium due to its mucin-degrading metabolism, as well as to other microorganisms with similar metabolic activities.

Certain *R. torques* strains have been found to be numerically dominant populations capable of degrading mucin oligosaccharides in the human colon due to their constitutive production of glycosidases [310]. *Bifidobacterium* species and *Clostridium perfringens* are also mucin-degrading bacteria found in the gut [310, 311], and certain strains of *B. fragilis* produce a wide range of hydrolytic enzymes that could also be involved in mucin degradation [312]. The mucus layer of the gastrointestinal mucosa has important physiological roles such as, for

example, enabling the stabilisation of some commensal bacteria and protecting from the association of pathogens. On the other hand, mucin degradation could serve as nutritional support for enteric bacteria, especially when other nutrients are absent, and could alter the glycolipids and glycoproteins of colonocyte membranes leading to alterations to the attachment sites for toxins and bacteria.

Therefore, an alteration in the mucus layer might have important consequences for the maintenance of a homeostatic situation. These alterations can be either quantitative (reduction of thickness) or qualitative (changes in mucin glycosylation). Although mucin-degrading bacteria are present at low levels in healthy humans (representing about 1% of total cultivable fecal bacteria), an alteration of the microbial composition characterised by the enrichment in mucolytic bacteria may occur under certain conditions. In a study that compared the ileal microbiota of piglets fed by total parenteral nutrition with that of others fed with total enteral nutrition, it was observed that the former had an increase in mucolytic bacteria driven by the lower nutrient availability [313]. A reduction of the mucus layer in IBD has been reported and is especially evident in UC patients [215]. In addition, increased levels of fecal mucinase and sulphatase activity, apparently of bacterial origin, were detected in UC fecal extracts [314, 315]. A genetic component in the host was also suspected of being involved in the production of certain mucins in UC patients [137]. In contrast, bacterial glycosidase activities, especially β-galactosidase (β-gal), were reduced in feces from patients with active CD [221, 316]. A possible role of the mucus layer in IBD pathogenesis seems to be more probable for UC patients than CD patients.

The modification of glycans might also have significant consequences to the host-bacteria relationship because novel adhesion sites for toxins and bacteria can appear [310]. It has been suggested that changes in mucin glycosylation, due to host-genetic alterations, may contribute to an increased association of bacteria within the mucus [317]. Conte *et al.* [220] supported this idea and added that changes in glycoprotein components of the intestinal mucosa could determine loss or unmasking of certain receptors leading to a different selection of microorganisms, for example decreasing *B. vulgatus* and favouring *E. coli*. An alteration in microbial composition, such as occurs in CD, could also explain this situation if certain microorganisms were responsible for the changes in enterocyte glycosylation.

In conclusion, further studies are needed to better address this question, focusing on mucosa-associated bacteria and aiming at: i) the quantification of total mucolytic bacteria in IBD patients and control subjects and ii) the identification

of putative specific changes caused to the glycans of enterocyte membranes and finding out which bacterial species are involved in such changes and what the consequences are to the host.

1.2 Crohn's colitis is microbiologically closer to ulcerative colitis than Crohn's ileitis

Crohn's disease can involve any part of the gastrointestinal tract, but the terminal ileum is most frequently involved – two thirds of CD patients – and a third of patients have isolated ileal disease [5]. Recent studies have provided several lines of evidence suggesting that Crohn's ileitis is microbiologically distinct from Crohn's colitis. This evidence includes:

i) Dramatic changes in the global gut microbiota were particularly evident for individuals with ileal CD as revealed by T-RFLP [180]. However, we should consider that, to some extent, the misclassification of Crohn's colitis and UC patients could lead to a homogenisation of their microbial composition, thus minimising the differences between both phenotypes. Occasionally, these IBDs are incorrectly diagnosed because of their similar clinical manifestations, so that doctors need to change the diagnosis of their patients as the disease develops.

ii) Increased abundance of *E. coli* has been repeatedly reported for I-CD patients, as evidenced by quantitative PCR [207, 219, 318]. In addition, we observed that this increment was irrespective of the sampled zone along the bowel, thus indicating that these differences were not from site-specific samples but from CD phenotype-specific samples, as others have subsequently corroborated [219].

iii) Decreased levels of *F. prausnitzii* were particularly significant in I-CD patients [219]. But the results were not conclusive in IC-CD patients, since two out of three patients turned out to be more similar to C-CD and one out three closer to I-CD patients.

iv) The observation that the AIEC pathovar was especially associated with I-CD [207, 288]. We later observed that the prevalence of AIEC was similar among I-CD, C-CD or IC-CD patients; however, the pathovar was more abundant and more diversified in CD patients with ileal localisation [318].

v) Reduced expression of α-defensins and reduced antimicrobial activity in I-CD but not in C-CD, UC or pouchitis [144]. The reduction in α-defensin might

have an influence on the entire bacterial community, by enhancing the multiplication of bacteria, especially those species which are the target of this antimicrobial peptide, thus favouring dysbiosis. A study of the effects of α-defensins on *E. coli* and *F. prausnitzii* populations would help to elucidate whether the alteration in α-defensin production is the primary cause of dysbiosis in I-CD.

In addition to the specific microbial imbalances, Crohn's ileitis is achieving a distinct identity with respect to other IBD types due to the positive correlation with several genetic risk factors ([319] and references therein):

i) *CARD15-NOD2* frameshift mutation is particularly associated with ileal-CD.

ii) There is a reduction of ileal α-defensin production in CD patients, which is in turn associated to *CARD15-NOD2* deficiencies.

iii) CEACAM6, which is the receptor for type 1 pili-mediated adhesion of AIEC, is overexpressed in the ileum of ileal-CD patients.

iv) The 300 A/A genotype of the autophagy-related gene *ATG16L1* conferred higher susceptibility to ileal disease than to CD of the colon.

Conversely, no specific bacterial species have been consistently described for Crohn's colitis, despite this group of patients showing an altered intestinal microbial community. Nevertheless, a lower HBD2 gene copy number in the β-defensin locus has been described as predisposing the subject to colonic CD, probably through diminished β-defensin expression, whereas it is strongly expressed in UC patients [151].

In light of these observations, future research should differentiate each phenotype of IBD to better describe the microbiological, environmental and genetic aspects associated with them. Although most recent studies detail the type of CD localisation, the results for CD are frequently analysed globally, which could lead to a blurring of the putative differences. Alterations of the microbial community structure in the gut have recently been described for other diseases, such as colorectal cancer and polyposis [303, 320, 321], obesity [322], irritable bowel syndrome [323, 324], celiac disease [325-328] and diabetes [329]. Given the sensitivity of the intestinal microbiota, non-IBD controls must also be well chosen, avoiding the inclusion of patients with these alterations in comparative studies.

1.3 Relevance and applicability of dysbiosis in Crohn's disease

1.3.1 Applicability to alternative diagnostic tools for inflammatory bowel diseases

The imbalances in the microbial community structure in IBDs can be used for diagnostic purposes [180, 205, 212, 215, 219]. Occasionally, CD (especially Crohn's colitis) and UC are incorrectly diagnosed due to their similar clinical manifestations, thus being misclassified or included within the category 'indeterminate colitis'. Distinguishing between these types of IBDs is necessary because important decisions depend on it, such as whether to use surgery or not, or the dosage, administration route and type of medication to be used. For this reason, new specific, reliable and sensitive tools that might help to diagnose the different types of IBDs would be of great interest.

Molecular-based tools represent an advance in the investigation of microbial populations from complex environments like the intestine. As the differences between IBDs are often differences in quantity rather than presence/absence of certain microorganisms, quantitative PCR is a promising tool for IBD diagnosis based on the microbial community structure. It is a reliable and robust technique that can be easily introduced in any laboratory and has the necessary specificity to quantify a range from particular species up to higher taxonomic levels. The most difficult step to overcome during the design of such a molecular diagnostic tool is to find suitable bacterial (disease-specific) indicators and design specific primers suitable for this task.

To date, the most consistent bacterial populations used to differentiate CD from UC are *F. prausnitzii* and *E. coli*. However, differences in their abundances appear to be useful for distinguishing I-CD from UC, but not C-CD from UC [219]. This is an important limitation since the difficulty in diagnosing CD from UC occurs mainly with C-CD patients. For this reason, an essential issue for gut microbiologists is to find additional biomarkers that are specific for C-CD and UC. It has been observed that *F. prausnitzii* abundances vary with disease activity in CD patients [215], making it an effective biomarker for active CD patients only. Nevertheless, the selection of unstable microbial populations as indicators would be advantageous, since these biomarkers could be used as predictors of a higher relapse-risk in IBD [295].

The fact that the mucosa-associated microbiota is stable along the digestive tract [208], between ulcerated and non-ulcerated mucosa [206] and between inflamed and non-inflamed areas [204] is an advantage because no specific sample site

needs to be selected. Nevertheless, the possibility of using a non-invasive sample, such as fecal samples, would definitely improve the applicability of the diagnostic tool.

Swidsinski *et al.* [215] have already designed an alternative diagnostic tool based on the quantification, using FISH, of *F. prausnitzii* and host leukocytes from fecal samples. They achieved a high specificity and acceptable sensibility. However, they did not separate CD patients according to disease localisation and they were only able to distinguish with high certainty between active CD patients and UC patients. Those patients with a remission period longer than one year could hardly be differentiated from UC patients. Moreover, some overlap was observed between CD and celiac disease. Their results show the difficulty in finding specific, disease-activity-independent indicators. Including new bacterial indicators could partially solve these limitations. On the other hand, using quantitative PCR rather than FISH might facilitate the analysis and could be more easily introduced into diagnostic laboratories.

1.3.2 Dysbiosis: cause or consequence?

Given the heterogeneity that characterises the CD-associated microbiota, the question of whether dysbiosis is a cause or a consequence of the disease remains unresolved [196]. Even defining whether the whole bacterial community is involved, or if the presence of certain commensal or opportunistic pathogens are responsible, or if simply a single pathogen is the cause, is still under debate. Those works involving newly diagnosed patients and/or patients in remission may help to shed light on this question, since putative changes in microbial composition due to medication or inflammation are diminished. However, we observed that neither the activity of the disease nor the type of medication could segregate CD patients according to their microbial composition.

Lupp *et al.* [330], using animal models, demonstrated that inflammation is sufficient to produce a dysbiosis favouring aerotolerant bacteria, such as *E. coli*. This suggests that dysbiosis in CD is a consequence of the inflammatory process. However, many studies based on IBD patients have produced results that go against this theory and suggest a causative role for microorganisms, at least for certain CD phenotypes. If inflammation was the cause of dysbiosis, then different phenotypes of IBD (ileal CD, colonic CD, and UC) would share certain bacterial compositional features, since they all involve inflammation. However, several studies have demonstrated that UC and CD, especially ileal CD, show particular microbial profiles [203, 207, 212, 219], which suggests that the microorganisms

play different roles in each IBD phenotype. Moreover, when the microbiota of inflamed and non-inflamed areas of CD-mucosa were compared, no differences were observed [203, 205, 212, 214, 217]. Unveiling which factors determine the patchy distribution of inflamed areas in CD is a difficult task. Nevertheless, the answer may lie in a scenario in which a minimum bacterial density is reached and where the GALT is more abundant.

Twin-based studies are helpful to understand the importance of microorganisms in CD since differences in the gut microbiota are independent of host genetics. Dicksved *et al.* [180] demonstrated that T-RFLP fingerprints were more similar between CD- concordant twins than between discordant twins, suggesting that the diseased individuals had a different microbial community structure from their healthy twins regardless of their genetic similitude. Another study from the same laboratory reported that, with regard to the abundance of certain specific bacterial populations (*F. prausnitzii* and *E. coli*), discordant I-CD twins could be segregated according to their disease phenotype rather than their twin pair. These results confirmed their previous observation that intestinal microbiota is associated with disease rather than host genotype.

It can be hypothesised that some species specifically associated with CD may have a causative role by somehow damaging the host (for example, AIEC), others may cause a loss of protection against pathogens (for example *F. prausnitzii*), and other microorganisms may only be a consequence of the environmental changes brought about by inflammation. Defining the specific role of each microorganism is an extremely difficult task in such a complex community; especially if we consider that synergistic relationships between microorganisms could also be noteworthy in CD pathogenesis. Defining the sequence of the processes chronologically is also complicated. Going back to the work of Lupp *et al.* [330], although *Enterobacteriaceae* could be a consequence of a primary inflammation, this would not rule out their putative implication in CD pathogenesis. These bacteria could be a subsequent part of the etiology by triggering chronic inflammation (considering that those clones better able to persist in the mucosa, intra- and/or extracellularly, would be selected) or by worsening the clinical symptoms (for example by granuloma formation).

2 Role of *Escherichia coli*, especially adherent-invasive *E. coli*, in Crohn's disease

A number of culture-based and molecular-based studies support the theory that *E. coli* is a microbiological factor implicated in CD. The higher abundance of *E. coli* in CD patients, in conjunction with the detection of this bacterium inside intestinal epithelial cells, mesenteric lymph nodes, and granulomas, point to a putative etiologic role rather than a consequence of inflammation. Intestinal microbiota may be involved in CD in two ways: i) dysbiosis of the commensal microbiota, in which protective bacteria decrease as harmful bacteria increase; ii) a low-grade infection by a persistent pathogen, either traditional or opportunistic. A putative role for *E. coli* in CD fits both hypotheses. Non-pathogenic *E. coli* are common colonisers of the mucus layer of the intestinal tract and have a mutualistic relationship with their hosts. They rarely cause disease, but in some special situations they can cause infections, for example, when the host is immunocompromised or when the intestinal barrier is broken. In turn, pathogenic *E. coli* have acquired virulence factors that allow them to actively cause several intestinal and extra-intestinal infections, in humans and animals. Finding the new potential AIEC pathovar associated with CD supports the theory of a low-grade infection by a persistent opportunistic pathogen, which is not mutually exclusive with dysbiosis, because the reduction of protective bacteria such as *F. prausnitzii* might contribute by worsening the situation. However, the precise role of *E. coli* in CD must still be clarified, including putative complementary host factors or host-bacterium / bacterium-bacterium interactions.

2.1 Evidence of AIEC implication in Crohn's disease

The features of the AIEC strain LF82 described to date have been comprehensively linked to many characteristics of CD pathogenesis, thus making this pathovar the subject of numerous studies. First, the AIEC adhesion to and invasion of intestinal epithelial cells induces inflammatory responses [277]. Second, AIEC are able to reduce barrier functions by disorganising F-actin and displacing ZO-1 and E-cadherin from the apical junctional complex [278]. Third, their ability to survive and replicate extensively in large vacuoles without inducing cell host death and promoting the secretion of TNF-α indicates that

AIEC may be involved in chronic antigenic stimulation and T-cell and macrophage activation [279], as well as in the formation of granulomas [280].

The implication of AIEC in CD is achieving increasing relevance since several independent studies have revealed a higher prevalence of this pathovar in CD patients [207, 278, 288, 289, 318]. In comparison with previous works an increase in the prevalence of AIEC was observed in our study [318]. It is probably due to the methodological approach used, which enabled us to obtain a more accurate prevalence value by analysing a greater number of *E. coli* per patient. In addition, we have described for the first time the relative abundance, richness and diversity of AIEC strains within the mucosa-associated *E. coli* population. Although the AIEC abundance was low and variable, it was higher in CD patients than it was in controls, especially in Crohn's ileitis patients. This could be partially explained by the increase in CEACAM6 receptors described for the ileal mucosa of CD patients [287]. In turn, the higher richness of AIEC subtypes found in CD patients leads us to hypothesise that this pathovar might be more permanent in CD intestinal mucosa, since a greater persistence in the environment could eventually contribute to the diversification of clones.

2.1.1 Adherent and invasive *E. coli* strains in cow mastitis, granulomatous colitis of Boxer dogs and human colorectal cancer

Similar adherent and invasive *E. coli* strains have also been associated with persistent bovine mastitis [331] and granulomatous colitis in Boxer dogs [332]. These are important findings for research into models for comparing the pathogenicity of adherent and invasive *E. coli* in IBD and other diseases, as well as research into putative zoonosis. *E. coli* strains associated with persistent bovine mastitis were found to be adherent, invasive and able to survive and replicate within the MAC-T bovine mammary epithelial cell line. However, the majority of them belonged to A phylogenetic groups, whereas CD-associated AIEC strains belong principally to B2 and D phylogroups. In addition, *E. coli* strains isolated from granulomatous colitis in Boxer dogs belonged mainly to the B2 phylogroup and shared a variety of genes implicated in iron acquisition and some other virulence factors characteristic of extraintestinal pathogenic *E. coli*, much like CD-associated AIEC strains [207, 318]. Moreover, they had adhesion, invasion and intra-macrophage replication properties similar to AIEC strain LF82, which further supports the link between intra-macrophage replication and granuloma formation.

Adhesive and invasive mucosa-associated *E. coli* have also been detected in colorectal cancer, especially beneath the colonic mucus layer. Martin *et al.* [289] proposed that colorectal cancer may be the result of a low-grade continuous inflammation and failure of apoptosis driven by the presence of these intracellular bacteria. Nevertheless, differences between isolates of *E. coli* from CD and colorectal patients in the distribution of UPEC virulence-associated genes and pathogenicity islands have been recently described by the same research group [333].

2.1.2 Host factors and AIEC

Although AIEC strains are found at low frequency in non-IBD controls, their presence in these subjects indicates that additional factors are needed to cause disease. Among them, CD-specific genetic susceptibility loci such as *CARD15*, *ATG16L1*, and *IRGM* could be involved in the handling of AIEC infection [290]. *CARD15* variants have been especially associated with ileal and stricturing disease phenotypes, as well as early disease onset. An alteration in the expression or operation of the *NOD2-CARD15* may have important effects in bacterial recognition, tolerance and clearance. Since AIEC are more prevalent and abundant in CD patients with ileal involvement [207, 288, 318], a possible relation between a defective innate immunity and invasive *E. coli* could explain the ileal phenotype of CD [334]. Consistent with this, a recent study observed that the *in vitro* infection of CD monocytes carrying *CARD15* polymorphisms with the AIEC strain LF82, resulted in a disturbed early inflammatory response, characterised by an early reduced IL-1β, IL-6 and IL-10 cytokines production, with respect to wild-type monocytes [335]. Moreover, *CARD15* 3020insC (1007fs) variant was reported to correlate with decreased α-defensin production in I-CD patients [146]. In a situation where α-defensins are depleted, the development of AIEC and its colonisation of the mucus layer (along with other intestinal bacterial populations) could be enhanced. As AIEC strains are intracellular pathogens, defects in autophagy-related genes *ATG16L1* and *IRGM* could predispose patients to CD by promoting prolonged survival of intracellular microorganisms within host cells. Given the possible link between host susceptibility and a dysfunctional immunological response to a persistent infection by intracellular pathogens and their clearance, further studies investigating the link between expression of *CARD15*, *ATG16L1*, or *IRGM* variants and AIEC infection will be of great interest [290]. In turn, the CEACAM6 receptor, site of AIEC adhesion by type 1 pili, has been proposed as a host susceptibility factor that could be involved in I-CD [287]. Nevertheless,

additional studies are necessary to confirm whether the overexpression of CEACAM6 in I-CD is a primary or secondary event of the inflammatory process.

The relevance of host factors in CD pathogenesis has recently been supported in a study that describes a defect in the follicle-associated epithelium of ileal CD patients that entails increased bacterial adhesion and uptake, both transcellular and intercellular, of non-pathogenic *E. coli* [336]. Following bacterial uptake, higher quantities of commensal *E. coli* then co-localised with dendritic cells in CD tissue, where levels of TNF-α were also augmented. Whether or not commensal *E. coli* strains are implicated in CD pathogenesis is an intriguing question. The work of Keita *et al*. emphasises the importance of host factors and plays the relevance of *E. coli* pathogenic features down. However, the authors point to the necessity of further studies using *E. coli* strains isolated from IBD patients. Hypothetically, a balance between the degree of virulence of the *E. coli* strains (for example, increased colonisation factors) and the degree of host susceptibility could determine the risk of contracting CD.

2.1.3 Synergism with other microorganisms?

Possible synergism between AIEC and other microorganisms is another factor to be considered, especially in such an environment as the intestine, where up to 10^{14} microorganisms share the same biotope. Recent studies have demonstrated that microbial mannans can suppress the mucosal phagocyte function causing dose-related increased survival of CD-associated *E. coli* HM605 within adherent monocytes [337]. *Mycobacterium avium paratuberculosis*, *Candida albicans* and *Saccharomyces cerevisiae* are some examples of microorganisms expressing microbial mannans, in particular, the ASCA epitope. Therefore, these microorganisms could be contributing to the process.

In addition, a number of reports have provided evidence of synergistic effects between *Bacteroides fragilis* and *E. coli* resulting in the inhibition of macrophage phagocytosis, augmented bacterial growth and abscess formation [338]. *B. fragilis* and *E. coli* are the principal colonisers of mucosal biofilms in CD patients [216]. The possible synergism between these two bacterial species and their putative role in CD pathogenesis could be an interesting subject for further studies.

2.2 Characterisation of the AIEC strain collection

AIEC strains isolated from our study were diverse in serotype, overall genotype and phylogenetic origin, as observed previously [207, 332]. They shared a combination of several virulence-associated genes characteristic of extraintestinal pathogenic *E. coli* and the phylogroups B2 and D were the most abundant. However, a similar distribution of these genes was also detected in non-AIEC strains associated with the intestinal mucosa. Virulence factors from extraintestinal pathogenic *E. coli* have already been described for the resident colonic microbiota in normal mucosa [339, 340]. It has been speculated that these factors are involved in complex host-commensal niche colonisation and bacterial fitness, with virulence arising as a by-product [341, 342].

It is worth noting that although AIEC and extraintestinal pathogenic *E. coli* show a similar frequency of phylogenetic groups and share many virulence factors, most of the extraintestinal pathogenic *E. coli* did not behave like AIEC. This suggests that AIEC strains comprise a particular group of *E. coli* that is closely related to extraintestinal pathogenic *E. coli*, but distinguishable by phenotypic traits. The genetic machinery responsible for their phenotype remains unclear.

In this work, we have addressed biofilm formation as a new phenotypic feature of AIEC that may be involved in their pathogenesis. While we could not find genetic differences between AIEC and non-AIEC strains isolated from the intestinal mucosa, we observed that AIEC strains had stronger biofilm formation abilities than non-AIEC strains. Something that we did not carry out, but which would merit further investigation, was to perform the biofilm formation assays *in vivo* and to elucidate whether or not biofilm formation of AIEC is involved in the pathogenesis of Crohn's disease. Frequently, non-classical pathogenic bacteria that form biofilms are responsible for chronic infections [343]. Biofilms confer better bacterial establishment, protection against phagocytosis and antimicrobial agents and a higher probability of genetic exchange. They also facilitate the colonisation of new environments by their natural detachment process, which allows the dispersal of high infective doses [344]. Some of these traits may be highly relevant within the course of CD (discussed in chapter 3.1).

2.2.1 AIEC pathogenic features

Intensive research looking for AIEC strain LF82-specific elements associated with pathogenesis using mutagenesis-complementation approaches, have resulted in the finding some implicated genes. However, these genes are also present in

the non-pathogenic *E. coli* strain K12 [276, 281-287]. Differences in gene expression, sometimes dependent on LF82 genetic background, may determine the differential adhesion, invasion and intra-macrophage replication capacities of the strains [284, 286]. The virulence genes detected in LF82 in other studies using PCR methods targeting specific virulence genes, genome subtraction, plasmid isolation and sequencing are summarised in Table 23.

Among the virulence factors of the AIEC strain LF82, type 1 pili are involved in the colonisation of host tissue, as happens with UPEC and DAEC pathovars. Type 1 pili mediate adherence to intestinal epithelial cells via the CEACAM6 receptor and are also involved in the invasive ability of the AIEC strain LF82 by inducing membrane extensions, which surround the bacteria at the sites of contact [276, 287]. All AIEC strains from our collection harboured the *fimH* gene, which encodes for the fimH apical adhesin of type 1 pili. Interestingly, *fimH* alleles of LF82 were identical to certain *E. coli* strains isolated from granulomatous colitis in Boxer dogs [332] and similar to UPEC CFT073, indicating that they share similar receptor affinities. We also obtained positive results for the avian-pathogenic variant of the FimA major pilus subunit, $fimAv_{MT78}$, in LF82 and 18.2% of AIEC strains from our collection (Table 23), adding evidence to the similarities existing between AIEC and extraintestinal pathogenic *E. coli*.

Flagella, which have a linked synthesis with type 1 pili [345], are also involved in the invasion of intestinal epithelial cells via motility and by downregulating type 1 pili expression [281]. Moreover, flagella can trigger mucosal inflammation via the TLR5 receptor [346]. We found that biofilm formation was also dependent on flagella, since non-motile strains from our collection and LF82-Δ*fliC* mutant were not able to form biofilms. The positive correlation between higher adhesion and invasion levels and stronger biofilm formation abilities leads us to postulate that the machinery implicated in achieving the "AIEC phenotype" may share some factors with the machinery necessary for biofilm formation, such as type 1 pili and flagella, and/or may be related to a coordinated expression.

Table 23. Virulence genes present in LF82 described to date (data from genome sequencing is not included) and the frequency of some of these genes in our AIEC collection (N = 22).

Gene	Gene function	LF82	Ref.	Prevalence in the AIEC strains (%)
Genes characteristic of diarrhoeagenic *E. coli*				
bfp	Involved in EPEC bundling-forming pilus formation	–	[277]	0
aggR	Transcriptional activator of EAggEC	–	[277]	
eae	Intimin which mediates intimate attachment and invasion of EPEC	–	[275]	0
tia	Outer membrane protein involved in ETEC invasion	–	[275]	
pCVD432	EaggEC plasmid involved in adhesion	–	[289]	0
ipaH	Invasion protein of *Shigella* and EIEC	–	[332]	0
ipaC	Invasion protein of *Shigella* and EIEC	–	[275]	
eltA	*E. coli* heat labile enterotoxin of ETEC	–	[332]	0
est	*E. coli* heat stable enterotoxin	–	[332]	0
SLT	Shiga-like toxin	–	[332]	
Stx1 and 2	Shiga toxin 1 and 2 of EHEC	–	[277]	0
Adhesins				
fimH	Encodes for type 1 fimbriae D-mannose-specific adhesin	+	[276]	100
$fimAv_{MT78}$	Avian-pathogenic variant of fimA major subunit of type 1 fimbriae	+	[276]	18.2
papC	Gene involved in pyelonephritis-associated (P) pilus formation. Pillus assembly.	–	[332]	54.5
papA	Gene involved in pyelonephritis-associated (P) pilus formation. Major structural subunit of P fimbrial shaft; defines F antigen.	–	[332]	
papGII	Gene involved in pyelonephritis-associated (P) pilus formation. Gal(α1-4)Gal-specific pilus tip adhesin molecule, allele II.	–	[332]	
sfa/focDE	Type S pili and type 1C fimbriae	–	[332]	18.2
afa/draBC	Afimbrial Dr-binding adhesins	–	[288]	13.6
flmA54	Flagellin (fliC variant)	–	[332]	
focA	Major fimbrial subunit of F1C fimbriae	–	[332]	
lpfA	Major subunit of long polar fimbriae LPF	+	[332]	
lpfA (O113)	Major subunit of long polar fimbriae LPF (variant O113)	–	[332]	
Toxins				
cnf 1 i 2	Cytotoxic necrotising factor	–	[332]	27.3
cdt	Cytolethal distending toxin	–	[332]	13.6
hlyA	α-Hemolysin	–	[332]	27.3
pic	Serine protease involved in colonisation	–	[332]	
ccdB	Cytotoxic protein (targets DNA gyrase)	–	[332]	
rtx	Putative RTX family exoprotein	–	[332]	
Siderophores – Iron transport				
iucD	Aerobactin siderophore	–	[332]	54.5
iutA	Aerobactin siderophore	–	[332]	
iroN	Siderophore receptor	–	[332]	
fepC	Ferric enterobactin transport ATP-binding protein	+	[332]	
Protectins				
iss	Increased serum survival and surface exclusion protein	+	[332]	
kpsMII	Protein involved in polysialic acid transport, group II	+	[332]	
neuC	K1 gene implicated in sialic acid synthesis (capsule formation)	–	[332]	13.6
traT	Complement resistance protein	–	[332]	
ColV plasmid	Involved in aerobactin production and complement resistance of APEC	+	[207]	
Miscellaneous				
ibeA	Brain microvascular endothelial cell invasion gene	+	[332]	13.6
malX	Maltose- and glucose-specific IIABC component, PAI associated	+	[332]	
usp	Uropathogenesis-specific protein	+	[332]	
cia	Colicin Ia structural protein	–	[332]	
colY	Colicin Y structural protein	–	[332]	
mchB	Microcin H47 structural protein	–	[332]	
chuA	Outer membrane receptor protein, heme utilisation/transport protein	+	[332]	

Table 23. Virulence genes present in LF82 described to date (*Continued*).

Genes characteristic of other pathogens			
irp1 and *2*	Yersiniabactin biosynthetic protein (putative ligase)	+	[332]
fyuA	Pesticin/yersiniabactin receptor protein	+	[332]
hcp	Hemolysin coregulated protein (*Vibrio*)	+	[207]
ratA	Non-fimbrial adhesin of *Salmonella typhimurium* that is associated with colonisation of the cecum and Peyer's patches	+	[207]
pMT1-like plasmid	Plasmid similar to *pMT1* of *Yersinia pestis*	+	[207]

Although not present in LF82, we detected other factors related to adhesion in a fraction of the AIEC collection. Among them, the UPEC-associated *papC* was present in more than 50% of AIEC strains, type S fimbriae in 18%, and Afa/Dr adhesins – characteristic of UPEC and DAEC strains – in 13% of the AIEC collection. Interestingly, AfaE adhesin interacts with the DAF epithelial receptor (also named CD55) triggering MICA (MHC class I-related) expression and further IFN-γ production. It has been suggested that this host-bacterium relationship plays a role in CD because higher levels of MICA expression were detected at the surface of colonocytes from CD patients compared with controls [347]. Moreover, some Afa/Dr adhesins interact with CEACAM receptors [348], which are known to be overexpressed in the ileum of Crohn's ileitis patients [287]. Therefore, the simultaneous presence of type 1 pili and Afa/Dr adhesins may confer advantages to enterocyte colonisation for those strains with both adhesion factors. Moreover, *ibeA* adhesin, which is involved in the invasion by the extraintestinal pathogenic *E. coli* and present in LF82, was found in only 14% of the AIEC strains. In addition to these colonisation factors, we also found toxins and molecules involved in iron acquisition and capsule formation. Chromosomal and episomal elements that are homologous to those described in UPEC, APEC and other intestinal pathogens such as *Salmonella* and *Yersinia* are present in LF82 and other AIEC strains [207, 332].

2.2.2 Comparative differential gene expression between AIEC and non-AIEC strains with identical PFGE profile: looking for the genes responsible for the "AIEC phenotype"

It is intriguing that extremely distant *E. coli* strains (different virulence gene profile, MLST and phylogenetic group) share the "AIEC phenotype"; whereas other very genetically close *E. coli* strains (identical PFGE profiles) can be classified as either AIEC or non-AIEC (Figure 23). This observation leads us to believe that what marks the "AIEC phenotype" is the differential expression of certain key genes that are widely distributed among all types of *E. coli*, or at least among ExPEC-like strains. Another possibility is that mobile genetic elements that have passed undetected by PFGE carry genes involved in AIEC phenotype

expression. The dispersion of these mobile genetic elements among all types of *E. coli* would determine the AIEC phenotype irrespective of the phylogenetic distance among the strains. Isolates that share identical PFGE profiles, serotype and virulence genes, but differ from their AIEC phenotype offer the possibility of performing comparative studies that could make it possible to find specific sequences for AIEC, or for genes which are up/down-regulated in comparison to non-AIEC strains. Bronowski *et al*. [333] have already performed subtraction libraries with an invasive *E. coli* strain isolated from colorectal cancer versus the UPEC strain 536 and the non pathogenic *E. coli* JM109. Baumgart *et al*. [207] also performed genome subtraction from AIEC strains 541-1 (O–:H18, B1), 541-15 (O21:H33, A), LF82 (O83:H1, B2) isolated from I-CD patients and *E. coli* MG1655. However, they compared strains that are in fact very different from each other. Given the high genetic variability among *E. coli*, a more targeted discrimination, searching for differences between genetically close strains that only differ on their AIEC phenotype, would probably reduce the number of differences and only those genes most involved in producing the AIEC phenotype would appear in the subtraction library. Subtraction libraries targeting differential expression of genes would, nevertheless, be the best choice.

At present, the definition of the AIEC pathovar is based mainly on its phenotypic characteristics. Finding out which genes are involved in this phenotype will be of great importance not only for our understanding of AIEC pathogenicity but because it will also simplify epidemiologic studies since molecular-based methods are fast and allow large numbers of samples to be processed. It would be easier to detect this pathovar in large groups of CD patients from various countries and identify AIEC strains in animals (which may act as reservoirs) and other environments. Nevertheless, the great diversity of genotypes that confer the AIEC pathotype makes this a very difficult task: we would need to study several AIEC strains at the same time rather than only LF82, since this strain is not representative of all the strains that belong to this pathovar.

3 Concluding remarks

Significant progress in the understanding of Crohn's disease-associated microbiota has been made in the last years. However, it is still impossible to answer which microorganisms might be directly involved and which is their way of pathogenesis. At present, even the way in which bacteria would be implicated in Crohn's disease etiopathogenesis is still controversial. The following three

theories are supported: the unidentified persistent pathogen theory, the excessive bacterial translocation theory, and the dysbiosis theory [224]. The complexity of the intestinal environment in conjunction with the high inter-subject variability hinders the clarification of questions posed by many scientists working in this field. Inter-laboratory studies reaching a consensus on the selection of patients and controls, as well as sampling procedures, and ensuring a detailed description of the characteristics of Crohn's disease patients (genetic susceptibility loci, disease localisation, medication, etc.) would improve the quality of microbiological investigations and would lead to faster and more upheld results. In addition, interdisciplinary studies (genetic, immunologic, microbiologic, clinical and epidemiologic) are needed given the multifactorial nature of Crohn's disease pathogenesis.

Alterations in the intestinal microbial structure is a characteristic of patients with Crohn's disease and other inflammatory and non-inflammatory disorders and conditions (for example, ulcerative colitis, irritable bowel syndrome, celiac disease, obesity or old age). This means therapeutic and dietary modification of the gut is one possible way of 'microbial restoration' and eventual clinical improvement for patients. Nevertheless, considerable effort is required to select and administer the therapeutic agents, and to recognise the putative secondary effects on the overall intestinal microbiota and the patient's health.

The role of *Escherichia coli*, especially the AIEC pathovar, in Crohn's disease is becoming widely accepted. Nevertheless, many aspects of its involvement are still not understood. Is it a primary or secondary event? What are the external factors associated with its pathogenesis – genetic susceptibility or putative synergistic relationships with other microorganisms or both? Regardless of its possible involvement in Crohn's disease pathogenesis, however, a better description of the AIEC pathovar is also of interest to microbiologists in order to better describe the genetic variability and phylogenetic relationships among *E. coli* strains.

Conclusions

1. CD patients present a microbial composition that is distinguishable from non-IBD controls. The latter show a more homogeneous intestinal microbiota (some strains being found consistently in all subjects) than that found among CD patients, who present a higher variability of species.
2. Although few patients suffering from UC were involved in the study, mucosa-associated bacteria of UC patients is distinguishable from that of CD patients and microbiologically closer to non-IBD subjects, which supports the theory that bacteria are playing different roles in CD and UC.
3. Microbial composition in the intestinal mucosa of CD patients does not correlate with disease severity, medication or other clinical aspects.
4. *Faecalibacterium prausnitzii* is a good bacterial indicator of appropriate healthy intestinal microbiota. Given the importance of this bacterium in gut physiology, the implications of its depletion in CD patients might be of relevance as a factor contributing to CD pathogenesis and merits further investigation.
5. *Enterobacteriaceae* (principally *Escherichia coli*) were more prevalent among CD patients than non-IBD controls. *Clostridium* spp. and *Ruminococcus torques* were also found to be associated with CD patients, as well as some opportunistic pathogenic γ-*Proteobacteria*. These bacterial species are, therefore, possible indicators of CD and could be implicated in the pathogenesis of this disease.
6. Dysbiosis observed in CD patients can be used to design a bacterial-based diagnostics tool based on the abundance of certain bacterial markers, such as *Faecalibacterium prausnitzii* and *Escherichia coli.*
7. *E. coli* abundance is higher in CD patients than in non-IBD controls, adding evidence of its implication in CD pathogenesis. However, only those patients with Crohn's ileitis can be differentiated from UC patients and non-IBD controls.
8. *E. coli* strains are host-specific even among CD patients. This rules out the hypothesis of a single clone or specific clonal group associated with CD.
9. An ecological imbalance exists among CD patients and controls in relation to *E. coli* abundance, but not to *E. coli* richness or diversity at a sub-species level.

10. Mucosa-associated *E. coli* from both CD patients and controls mainly belong to B2 and D phylogenetic groups and share virulence features of extraintestinal pathogenic *E. coli* at similar frequencies.

11. Although *E. coli* strains from CD patients and controls are genetically similar, some phenotypic features related to virulence, such as the adhesion and invasion abilities and the survival and intra-macrophage replication capacity of the strains (characteristic of AIEC pathovar), indicate that *E. coli* populations differ in their pathogenic behaviour.

12. The prevalence of AIEC is higher in CD patients than in controls, in both the ileum and the colon. The prevalence values observed here are higher than those reported so far.

13. AIEC abundance (calculated as the proportion of AIEC over the whole *E. coli* population) and AIEC subtype richness (calculated as the number of distinct AIEC subtypes per patient) are higher in CD patients. This new ecological data on the AIEC pathovar suggests that this bacterium is well-established in CD-mucosa and is evidence of the putative role of AIEC in Crohn's disease.

14. Although the AIEC pathovar is found less frequently and in lower abundance in controls than in CD patients, their presence in non-IBD subjects suggests that host genetic or immunologic factors are necessary to contract Crohn's disease.

15. The AIEC pathovar appears to be a diverse group because it is composed of a variety of pulsotypes*, serotypes, genotypes and phylogenetic groups.

16. Producing the "AIEC phenotype" may involve either genetic elements that cannot be resolved by PFGE or the differential expression of certain key genes.

17. Biofilm formation is a novel feature associated with the AIEC pathovar. This feature may contribute to AIEC pathogenesis by allowing a more efficient colonisation of the intestinal mucosa, either extracellularly or intracellularly.

18. Despite AIEC and extraintestinal pathogenic *E. coli* (ExPEC) pathovars share similar virulence genes sets and certain strains are phylogenetically related, the majority of ExPEC do not behave as AIEC, thus confirming that AIEC pathovar is close to ExPEC but has virulence-specific features.

References‡

‡ Include all references cited in the manuscript, with the exception of those specifically referenced in Chapters 1-3.

1. Mendoza Hernández JL, Lana Soto R, Díaz-Rubio M: **Definiciones y manifestaciones clínicas generales.** In: *Enfermedad Inflamatoria Intestinal.* Edited by Gasull MA, Gomollón F, Hinojosa J, Obrador A, Third edn. Madrid: Arán Ediciones, S.L.; 2007: 21-28.

2. Hendrickson BA, Gokhale R, Cho JH: **Clinical Aspects and Pathophysiology of Inflammatory Bowel Disease**. *Clin Microbiol Rev* 2002, **15**(1):79-94.

3. Silverberg MS, Satsangi J, Ahmad T, Arnott ID, Bernstein CN, Brant SR, Caprilli R, Colombel JF, Gasche C, Geboes K, Jewell DP, Karban A, Loftus EV, Peña AS, Riddell RH, Sachar DB, Schreiber S, Steinhart AH, Targan SR, Vermeire S, Warren BF: **Toward an integrated clinical, molecular and serological classification of inflammatory bowel disease: Report of a Working Party of the 2005 Montreal World Congress of Gastroenterology.** *Can J Gastroenterol* 2005, **19**(Suppl A):5-36.

4. Louis E, Collard A, Oger AF, Degroote E, El Yafi FAN, Belaiche J: **Behaviour of Crohn's disease according to the Vienna classification: changing pattern over the course of the disease**. *Gut* 2001, **49**(6):777-782.

5. Gasche C, Scholmerich J, Brynskov J, D'Haens G, Hanauer SB, Irvine EJ, Jewell DP, Rachmilewitz D, Sachar DB, Sandborn WJ, Sutherland LR: **A simple classification of Crohn's disease: report of the Working Party for the World Congresses of Gastroenterology, Vienna 1998.** *Inflamm Bowel Dis* 2000, **6**(1):8-15.

6. Riera-Oliver J: **Definiciones conceptuales de la enfermedad inflamatoria intestinal. Concepto de cronicidad.** In: *Enfermedad Inflamatoria Intestinal.* Edited by Gassull M, Gomollón F, Hinojosa J, Obrador A, Second edn. Madrid: Ediciones Ergón; 2002: 3-5.

7. Vucelic B: **Colitis indeterminada**. In: *Enfermedad Inflamatoria Intestinal.* Edited by Gasull MA, Gomollón F, Hinojosa J, Obrador A, Third edn. Madrid: Arán Ediciones, S.L; 2007: 407-414.

8. Álvarez L: **Colitis indeterminada**. *Gastr Latinoam* 2007, **18**(2):228-230.

9. Evans PE, Pardi DS: **Extraintestinal manifestations of inflammatory bowel disease: focus on the musculoskeletal, dermatologic, and ocular manifestations.** *Med Gen Med* 2007, **9**(1):55.

10. Lakatos L, Pandur T, David G, Balogh Z, Kuronya P, Tollas A, Lakatos P: **Association of extraintestinal manifestations of inflammatory bowel disease in a province of western Hungary with disease phenotype: results of a 25-year follow-up study.** *World J Gastroenterol* 2003, **9**(10):2300-2307.

11. Salvarani C, Vlachonikolis IG, van der Heijde DM, Fornaciari G, Macchioni P, Beltrami M, Olivieri I, Di Gennaro F, Politi P, Stockbrügger RW, Russel MG: **Musculoskeletal manifestations in a population-based cohort of inflammatory bowel disease patients.** *Scand J Gastroenterol* 2001, **36**(12):1307-1313.

12. González-Huix F: **Enfermedades asociadas a la enfermedad inflamatoria intestinal.** In: *Enfermedad Inflamatoria Intestinal.* Edited by Gasull MA, Gomollón F, Hinojosa J, Obrador A, Third edn. Madrid: Arán Ediciones, S.L.; 2007: 423-432.

13. Higgins PDR, Davis KJ, Laine L: **The epidemiology of ischaemic colitis**. *Aliment Pharmacol Ther* 2004, **19**(7):729-738.

14. Agrawal A, Whorwell PJ: **Irritable bowel syndrome: diagnosis and management**. *BMJ* 2006, **332**(7536):280-283.

15. Eriksson ME, Andrén KI, Eriksson HT, Kurlberg GK: **Irritable bowel syndrome subtypes differ in body awareness, psychological symptoms and biochemical stress markers**. *World J Gastroenterol* 2008, **14**(31):4889-4896.

16. Lacy BE, Lee RD: **Irritable bowel syndrome: a syndrome in evolution.** *J Clin Gastroenterol* 2005 **39**(5 Suppl 3):S230-242.

17. Drossman DA, Morris CB, Hu Y, Toner BB, Diamant N, Leserman J, Shetzline M, Dalton C, Bangdiwala SI: **A prospective assessment of bowel habit in irritable bowel syndrome in women: Defining an alternator**. *Gastroenterology* 2005, **128**(3):580-589.

18. Ivison SM, Steiner TS: **Anti-flagellin antibodies in irritable bowel syndrome: another attack on our commensals?** *Neurogastroenterol Motil* 2008, **20**(10):1081-1085.

19. Drossman D: **What does the future hold for irritable bowel syndrome and the functional gastrointestinal disorders?** *J Clin Gastroenterol* 2005, **39**(5 Suppl 3):S251-256.

20. Olbe L: **Concept of Crohn's disease being conditioned by four main components, and irritable bowel syndrome being an incomplete Crohn's disease**. *Scand J Gastroenterol* 2008, **43**(2):234 - 241.

21. Sawczenko A, Sandhu BK: **Presenting features of inflammatory bowel disease in Great Britain and Ireland**. *Arch Dis Child* 2003, **88**(11):995-1000.

22. Domènech-Morral E, Mañosa-Círia M: **Clínica y criterios diagnósticos de la enfermedad de Crohn**. In: *Enfermedad Inflamatoria Intestinal.* Edited by Gasull MA, Gomollón F, Hinojosa J, Obrador A, Third edn. Madrid: Arán Ediciones, S.L.; 2007: 333-344.

23. Stange EF, Travis SPL, Vermeire S, Beglinger C, Kupcinkas L, Geboes K, Barakauskiene A, Villanacci V, Von Herbay A, Warren BF, Gasche C, Tilg H, Schreiber SW, Scholmerich J, Reinisch W: **European evidence based consensus on the diagnosis and management of Crohn's disease: definitions and diagnosis**. *Gut* 2006, **55**(suppl 1):i16-i35.

24. Geboes K: **Anatomía patológica de la enfermedad inflamatoria intestinal.** In: *Enfermedad Inflamatoria Intestinal.* Edited by Gasull MA, Gomollón F, Hinojosa J, Obrador A, Third edn. Madrid: Arán Ediciones, S.L.; 2007: 663.

25. Gisbert JP: **Marcadores biológicos y pronósticos**. In: *Enfermedad Inflamatoria Intestinal.* Edited by Gasull MA, Gomollón F, Hinojosa J, Obrador A, Third edn. Madrid: Arán Ediciones, S.L.; 2007: 663.

26. Benkhadra F, Humbel RL: **Les marqueurs sérologiques des maladies inflammatoires chroniques intestinales (MICI)**. *Immuno-analyse & Biologie Spécialisée* 2008, **23**(4):202-211.

27. Xavier RJ, Podolsky DK: **Unravelling the pathogenesis of inflammatory bowel disease**. *Nature* 2007, **448**(7152):427-434.

28. Sartor R: **Mechanisms of Disease: pathogenesis of Crohn's disease and ulcerative colitis**. *Nat Clin Pract Gastroenterol Hepatol* 2006, **3**(7):390-407.

29. Hanauer SB: **Inflammatory Bowel Disease: Epidemiology, Pathogenesis, and Therapeutic Opportunities**. *Inflamm Bowel Dis* 2006, **12**:S3-S9.

30. Aladrén BS, Gasull MA, Gomollón F: **Epidemiología de las enfermedades inflamatorias intestinales. Factores ambientales internos y externos en su patogenia.** In: *Enfermedad Inflamatoria Intestinal.* Edited by Gasull MA,

Gomollón F, Hinojosa J, Obrador A, Third edn. Madrid: Arán Ediciones; 2007: 29-50.

31. Karlinger K, Györke T, Makö E, Mester Á, Tarján Z: **The epidemiology and the pathogenesis of inflammatory bowel disease**. *Eur J Radiol* 2000, **35**(3):154-167.

32. Gent AE, Hellier MD, Grace RH, Swarbrick ET, Coggon D: **Inflammatory bowel disease and domestic hygiene in infancy**. *Lancet* 1994, **343**(8900):766-767.

33. Kruiningen HJV, Joossens M, Vermeire S, Joossens S, Debeugny S, Gower-Rousseau C, Cortot A, Colombel J-F, Rutgeerts P, Vlietinck R: **Environmental factors in familial Crohn's disease in Belgium**. *Inflamm Bowel Dis* 2005, **11**(4):360-365.

34. Hugot J-P, Alberti C, Berrebi D, Bingen E, Cézard J-P: **Crohn's disease: the cold chain hypothesis**. *Lancet* 2003, **362**(9400):2012-2015.

35. Card T, Logan RFA, Rodrigues LC, Wheeler JG: **Antibiotic use and the development of Crohn's disease**. *Gut* 2004, **53**(2):246-250.

36. Baumgart DC, Sandborn WJ: **Inflammatory bowel disease: clinical aspects and established and evolving therapies**. *Lancet* 2007, **369**(9573):1641-1657.

37. Orholm M BV, Sørensen TI, Rasmussen LP, Kyvik KO.: **Concordance of inflammatory bowel disease among Danish twins. Results of a nationwide study.** *Scand J Gastroenterol* 2000 **35**(10):1075-1081.

38. Thompson NP, Driscoll R, Pounder RE, Wakefield AJ: **Genetics versus environment in inflammatory bowel disease: results of a British twin study**. *BMJ* 1996, **312**(7023):95-96.

39. Tysk C, Lindberg E, Jarnerot G, Floderus-Myrhed B: **Ulcerative colitis and Crohn's disease in an unselected population of monozygotic and dizygotic twins. A study of heritability and the influence of smoking**. *Gut* 1988, **29**(7):990-996.

40. Binder V: **Genetic Epidemiology in Inflammatory Bowel Disease.** *Dig Dis* 1998, **16**(6):351-355

41. Orholm M, Munkholm P, Langholz E, Nielsen OH, Sorensen IA, Binder V: **Familial occurrence of inflammatory bowel disease**. *N Engl J Med* 1991, **324**(2):84-88.

42. Peeters M, Nevens H, Baert F, Hiele M, de Meyer AM, Vlietinck R, Rutgeerts P: **Familial aggregation in Crohn's disease: Increased age-adjusted risk and concordance in clinical characteristics**. *Gastroenterology* 1996, **111**(3):597-603.

43. 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**(3):573-579.

44. Baumgart DC, Carding SR: **Inflammatory bowel disease: cause and immunobiology**. *Lancet* 2007, **369**(9573):1627-1640.

45. Goyette P, Labbé C, Trinh TT, Xavier RJ, Rioux JD: **Molecular pathogenesis of inflammatory bowel disease: Genotypes, phenotypes and personalized medicine**. *Ann Med* 2007, **39**(3):177 - 199.

46. Fisher SA, Tremelling M, Anderson CA, Gwilliam R, Bumpstead S, Prescott NJ, Nimmo ER, Massey D, Berzuini C, Johnson C, Barrett JC, Cummings FR, Drummond H, Lees CW, Onnie CM, Hanson CE, Blaszczyk K, Inouye M, Ewels P, Ravindrarajah R, Keniry A, Hunt S, Carter M, Watkins N, Ouwehand W, Lewis

CM, Cardon L, Lobo A, Forbes A, Sanderson J, Jewell DP, Mansfield JC, Deloukas P, Mathew CG, Parkes M, Satsangi J: **Genetic determinants of ulcerative colitis include the *ECM1* locus and five loci implicated in Crohn's disease**. *Nat Genet* 2008, **40**(6):710-712.

47. Cho JH: **The genetics and immunopathogenesis of inflammatory bowel disease**. *Nat Rev Immunol* 2008, **8**(6):458-466.

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

49. Ogura Y, Bonen DK, Inohara N, Nicolae DL, Chen FF, Ramos R, Britton H, Moran T, Karaliuskas R, Duerr RH, Achkar J-P, Brant SR, Bayless TM, Kirschner BS, Hanauer SB, Nunez G, Cho JH: **A frameshift mutation in *NOD2* associated with susceptibility to Crohn's disease**. *Nature* 2001, **411**(6837):603-606.

50. Lakatos PL, Lakatos L, Szalay F, Willheim-Polli C, Osterreicher C, Tulassay Z, Molnar T, Reinisch W, Papp J, Mozsik G, Ferenci P: **Toll-like receptor 4 and NOD2/CARD15 mutations in Hungarian patients with Crohn's disease: phenotype-genotype correlations.** *World J Gastroenterol* 2005, **11**(10):1489-1495.

51. Lesage S, Zouali H, Cézard J-P, Colombel J-F, Belaiche J, Almer S, Tysk C, O'Morain C, Gassull M, Binder V, Finkel Y, Modigliani R, Gower-Rousseau C, Macry J, Merlin F, Chamaillard M, Jannot A-S, Thomas G, Hugot J-P: ***CARD15*/*NOD2* mutational analysis and genotype-phenotype correlation in 612 patients with inflammatory bowel disease**. *Am J Hum Genet* 2002, **70**(4):845-857.

52. Hampe J, Franke A, Rosenstiel P, Till A, Teuber M, Huse K, Albrecht M, Mayr G, De La Vega F, Briggs J, Günther S, Prescott NJ, Onnie CM, Häsler R, Sipos B, Fölsch UR, Lengauer T, Platzer M, Mathew CG, Krawczak M, Schreiber S: **A genome-wide association scan of nonsynonymous SNPs identifies a susceptibility variant for Crohn disease in *ATG16L1*.** *Nat Genet* 2007, **39**(2):207-211.

53. McCarroll SA, Huett A, Kuballa P, Chilewski SD, Landry A, Goyette P, Zody MC, Hall JL, Brant SR, Cho JH, Duerr RH, Silverberg MS, Taylor KD, Rioux JD, Altshuler D, Daly MJ, Xavier RJ: **Deletion polymorphism upstream of *IRGM* associated with altered *IRGM* expression and Crohn's disease**. *Nat Genet* 2008, **40**(9):1107-1112.

54. Parkes M, Barrett JC, Prescott NJ, Tremelling M, Anderson CA, Fisher SA, Roberts RG, Nimmo ER, Cummings FR, Soars D, Drummond H, Lees CW, Khawaja SA, Bagnall R, Burke DA, Todhunter CE, Ahmad T, Onnie CM, McArdle W, Strachan D, Bethel G, Bryan C, Lewis CM, Deloukas P, Forbes A, Sanderson J, Jewell DP, Satsangi J, Mansfield JC, Cardon L, Mathew CG: **Sequence variants in the autophagy gene *IRGM* and multiple other replicating loci contribute to Crohn's disease susceptibility**. *Nat Genet* 2007, **39**(7):830-832.

55. Economou M, Trikalinos TA, Loizou KT, Tsianos EV, Ioannidis JPA: **Differential effects of NOD2 variants on Crohn's disease risk and phenotype in diverse populations: A Metaanalysis**. *Am J Gastroenterol* 2004, **99**(12):2393-2404.

56. McGovern DPB, Hysi P, Ahmad T, van Heel DA, Moffatt MF, Carey A, Cookson WOC, Jewell DP: **Association between a complex insertion/deletion**

polymorphism in *NOD1* (*CARD4*) and susceptibility to inflammatory bowel disease. *Hum Mol Genet* 2005, **14**(10):1245-1250.

57. Özen SC, Dağlı Ü, Kiliç MY, Törüner M, Çelik Y, Özkan M, Soykan I, Çetinkaya H, Ülker A, Özden A, Bozdayı AM: ***NOD2*/*CARD15*, *NOD1*/*CARD4*, and *ICAM-1* gene polymorphisms in Turkish patients with inflammatory bowel disease**. *J Gastroenterol* 2006, **41**(4):304-310.

58. Zouali H, Lesage S, Merlin F, Cezard JP, Colombel JF, Belaiche J, Almer S, Tysk C, O'Morain C, Gassull M, Christensen S, Finkel Y, Modigliani R, Gower-Rousseau C, Macry J, Chamaillard M, Thomas G, Hugot JP: **CARD4/NOD1 is not involved in inflammatory bowel disease**. *Gut* 2003, **52**(1):71-74.

59. Rousseaux C, Lefebvre B, Dubuquoy L, Lefebvre P, Romano O, Auwerx J, Metzger D, Wahli W, Desvergne B, Naccari GC, Chavatte P, Farce A, Bulois P, Cortot A, Colombel JF, Desreumaux P: **Intestinal antiinflammatory effect of 5-aminosalicylic acid is dependent on peroxisome proliferator-activated receptor-{gamma}**. *J Exp Med* 2005, **201**(8):1205-1215.

60. Sugawara K, Olson TS, Moskaluk CA, Stevens BK, Hoang S, Kozaiwa K, Cominelli F, Ley KF, McDuffie M: **Linkage to peroxisome proliferator-activated receptor-[gamma] in SAMP1/YitFc mice and in human Crohn's disease**. *Gastroenterology* 2005, **128**(2):351-360.

61. Dubuquoy L, Dharancy S, Nutten S, Pettersson S, Auwerx J, Desreumaux P: **Role of peroxisome proliferator-activated receptor [gamma] and retinoid X receptor heterodimer in hepatogastroenterological diseases**. *Lancet* 2002, **360**(9343):1410-1418.

62. Brand S, Staudinger T, Schnitzler F, Pfennig S, Hofbauer K, Dambacher J, Seiderer J, Tillack C, Konrad A, Crispin A, Göke B, Lohse P, Ochsenkühn T: **The role of Toll-like receptor 4 Asp299Gly and Thr399Ile polymorphisms and *CARD15*/*NOD2* mutations in the susceptibility and phenotype of Crohn's disease.** *Inflamm Bowel Dis* 2005, **11**(7):645-652.

63. Figueroa C, Peralta A, Herrera L, Castro P, Gutiérrez A, Valenzuela J, Aguillón J, Quera R, Hermoso M: **NOD2/CARD15 and Toll-like 4 receptor gene polymorphism in Chilean patients with inflammatory bowel disease.** *Eur Cytokine Netw* 2006, **17**(2):125-130.

64. Franchimont D, Vermeire S, El Housni H, Pierik M, Van Steen K, Gustot T, Quertinmont E, Abramowicz M, Van Gossum A, Devière J, Rutgeerts P: **Deficient host-bacteria interactions in inflammatory bowel disease? The toll-like receptor (TLR)-4 Asp299gly polymorphism is associated with Crohn's disease and ulcerative colitis**. *Gut* 2004, **53**(7):987-992.

65. Gazouli M, Mantzaris G, Kotsinas A, Zacharatos P, Papalambros E, Archimandritis A, Ikonomopoulos J, Gorgoulis V: **Association between polymorphisms in the Toll-like receptor 4, CD14, and *CARD15*/*NOD2* and inflammatory bowel disease in the Greek population.** *World J Gastroenterol* 2005, **11**(5):681-685.

66. Oostenbrug LE, Drenth JPH, de Jong DJ, Nolte IM, Oosterom E, Dullemen HMv, van der Linde K, te Meerman GJ, van der Steege G, Kleibeuker JH, Jansen PLM: **Association between toll-like receptor 4 and inflammatory bowel disease**. *Inflamm Bowel Dis* 2005, **11**(6):567-575.

67. Török H-P, Glas J, Tonenchi L, Mussack T, Folwaczny C: **Polymorphisms of the lipopolysaccharide-signaling complex in inflammatory bowel disease: association of a mutation in the Toll-like receptor 4 gene with ulcerative colitis**. *Clin Immunol* 2004, **112**(1):85-91.

68. Borm MEA, Bodegraven AA, Mulder CJJ, Bouma GKG: **A NFKB1 promoter polymorphism is involved in susceptibility to ulcerative colitis**. *Int J Immunogenet* 2005, **32**(6):401-405.

69. Karban AS, Okazaki T, Panhuysen CIM, Gallegos T, Potter JJ, Bailey-Wilson JE, Silverberg MS, Duerr RH, Cho JH, Gregersen PK, Wu Y, Achkar J-P, Dassopoulos T, Mezey E, Bayless TM, Nouvet FJ, Brant SR: **Functional annotation of a novel NFKB1 promoter polymorphism that increases risk for ulcerative colitis**. *Hum Mol Genet* 2004, **13**(1):35-45.

70. Klein W, Tromm A, Griga T, Fricke H, Folwaczny C, Hocke M, Eitner K, Marx M, Duerig N, Epplen J: **A polymorphism in the CD14 gene is associated with Crohn disease.** *Scand J Gastroenterol* 2002, **37**(2):189-191.

71. Obana N, Takahashi S, Kinouchi Y, Negoro K, Takagi S, Hiwatashi N, Shimosegawa T: **Ulcerative colitis is associated with a promoter polymorphism of lipopolysaccharide receptor gene, CD14.** *Scand J Gastroenterol* 2002 **37**(6):699-704.

72. Peltekova V, Wintle R, Rubin L, Amos C, Huang Q, Gu X, Newman B, Van Oene M, Cescon D, Greenberg G, Griffiths A, St George-Hyslop P, Siminovitch K: **Functional variants of OCTN cation transporter genes are associated with Crohn disease**. *Nat Genet* 2004, **36**(5):471-475.

73. Rioux J, Daly M, Silverberg M, Lindblad K, Steinhart H, Cohen Z, Delmonte T, Kocher K, Miller K, Guschwan S, Kulbokas E, O'Leary S, Winchester E, Dewar K, Green T, Stone V, Chow C, Cohen A, Langelier D, Lapointe G, Gaudet D, Faith J, Branco N, Bull S, McLeod R, Griffiths A, Bitton A, Greenberg G, Lander E, Siminovitch K, Hudson T: **Genetic variation in the 5q31 cytokine gene cluster confers susceptibility to Crohn disease.** *Nat Genet* 2001, **29**(2):223-228.

74. van Bodegraven AA, Curley CR, Hunt KA, Monsuur AJ, Linskens RK, Onnie CM, Crusius JBA, Annese V, Latiano A, Silverberg MS, Bitton A, Fisher SA, Steinhart AH, Forbes A, Sanderson J, Prescott NJ, Strachan DP, Playford RJ, Mathew CG, Wijmenga C, Daly MJ, Rioux JD, van Heel DA: **Genetic variation in myosin IXB is associated with ulcerative colitis**. *Gastroenterology* 2006, **131**(6):1768-1774.

75. Annese V, Valvano M, Palmieri O, Latiano A, Bossa F, Andriulli A: **Multidrug resistance 1 gene in inflammatory bowel disease: a meta-analysis.** *World J Gastroenterol* 2006, **12**(23):3636-3644.

76. Brant S, Panhuysen C, Nicolae D, Reddy D, Bonen D, Karaliukas R, Zhang L, Swanson E, Datta L, Moran T, Ravenhill G, Duerr R, Achkar J, Karban A, Cho J: **MDR1 Ala893 polymorphism is associated with inflammatory bowel disease.** *Am J Hum Genet* 2003 **73**(6):1282-1292.

77. Stoll M, Corneliussen B, Costello C, Waetzig G, Mellgard B, Koch W, Rosenstiel P, Albrecht M, Croucher P, Seegert D, Nikolaus S, Hampe J, Lengauer T, Pierrou S, Foelsch U, Mathew C, Lagerstrom-Fermer M, Schreiber S: **Genetic variation in DLG5 is associated with inflammatory bowel disease.** *Nat Genet* 2004, **36**(5):476-480.

78. Libioulle C, Louis E, Hansoul S, Sandor C, Farnir F, Franchimont D, Vermeire S, Dewit O, de Vos M, Dixon A, Demarche B, Gut I, Heath S, Foglio M, Liang L, Laukens D, Mni M, Zelenika D, Gossum A, Rutgeerts P, Belaiche J, Lathrop M, Georges M: **Novel Crohn disease locus identified by genome-wide association maps to a gene desert on 5p13.1 and modulates expression of *PTGER4***. *PLoS Genetics* 2007, **3**(4):e58.

79. Negoro K, Kinouchi Y, Hiwatashi N, Takahashi S, Takagi S, Satoh J, Shimosegawa T, Toyota T: **Crohn's disease is associated with novel**

polymorphisms in the 5'-flanking region of the tumor necrosis factor gene. *Gastroenterology* 1999, **117**(5):1062-1068.

80. Song Y, Wu K, Zhang L, Hao Z, Li H, Zhang L, Qiao T, Li C, Fan D: **Correlation between a gene polymorphism of tumor necrosis factor and inflammatory bowel disease**. *Chin J Dig Dis* 2005, **6**(4):170-174.

81. Sýkora J, Subrt I, Dìdek P, Siala K, Schwarz J, Machalová V, Varvarovská J, Pazdiora P, Pozler O, F. S: **Cytokine tumor necrosis factor-alpha A promoter gene polymorphism at position -308 G-->A and pediatric inflammatory bowel disease: implications in ulcerative colitis and Crohn's disease.** *J Pediatr Gastroenterol Nutr* 2006, **42**(5):479-487.

82. Duerr RH, Taylor KD, Brant SR, Rioux JD, Silverberg MS, Daly MJ, Steinhart AH, Abraham C, Regueiro M, Griffiths A, Dassopoulos T, Bitton A, Yang H, Targan S, Datta LW, Kistner EO, Schumm LP, Lee AT, Gregersen PK, Barmada MM, Rotter JI, Nicolae DL, Cho JH: **A genome-wide association study identifies *IL23R* as an inflammatory bowel disease gene**. *Science* 2006, **314**(5804):1461-1463.

83. Stokkers P, Reitsma P, Tytgat G, van Deventer SJH: **HLA-DR and -DQ phenotypes in inflammatory bowel disease: a meta-analysis**. *Gut* 1999 **45**(3):395-401.

84. Yap LM, Ahmad T, Jewell DP: **The contribution of HLA genes to IBD susceptibility and phenotype**. *Best Pract Res Clin Gastroenterol* 2004, **18**(3):577-596.

85. Rioux JD, Xavier RJ, Taylor KD, Silverberg MS, Goyette P, Huett A, Green T, Kuballa P, Barmada MM, Datta LW, Shugart YY, Griffiths AM, Targan SR, Ippoliti AF, Bernard E-J, Mei L, Nicolae DL, Regueiro M, Schumm LP, Steinhart AH, Rotter JI, Duerr RH, Cho JH, Daly MJ, Brant SR: **Genome-wide association study identifies new susceptibility loci for Crohn disease and implicates autophagy in disease pathogenesis**. *Nat Genet* 2007, **39**(5):596-604.

86. Rajilic-Stojanovic M, Smidt H, de Vos VM: **Diversity of the human gastrointestinal tract microbiota revisited**. *Environ Microbiol* 2007, **9**(9):2125-2136.

87. Zoetendal EG, Akkermans ADL, De Vos WM: **Temperature gradient gel electrophoresis analysis of 16S rRNA from human fecal samples reveals stable and host-specific communities of active bacteria**. *Appl Environ Microbiol* 1998, **64**(10):3854-3859.

88. Suau A, Bonnet R, Sutren M, Godon J-J, Gibson GR, Collins MD, Dore J: **Direct analysis of genes encoding 16S rRNA from complex communities reveals many novel molecular species within the human gut**. *Appl Environ Microbiol* 1999, **65**(11):4799-4807.

89. Marteau P, Pochart P, Dore J, Bera-Maillet C, Bernalier A, Corthier G: **Comparative study of bacterial groups within the human cecal and fecal microbiota**. *Appl Environ Microbiol* 2001, **67**(10):4939-4942.

90. Zoetendal EG, von Wright A, Vilpponen-Salmela T, Ben-Amor K, Akkermans ADL, de Vos WM: **Mucosa-associated bacteria in the human gastrointestinal tract are uniformly distributed along the colon and differ from the community recovered from feces**. *Appl Environ Microbiol* 2002, **68**(7):3401-3407.

91. Hold GL, Pryde SE, Russell VJ, Furrie E, Flint HJ: **Assessment of microbial diversity in human colonic samples by 16S rDNA sequence analysis**. *FEMS Microbiol Ecol* 2002, **39**(1):33-39.

92. Eckburg PB, Bik EM, Bernstein CN, Purdom E, Dethlefsen L, Sargent M, Gill SR, Nelson KE, Relman DA: **Diversity of the human intestinal microbial flora**. *Science* 2005, **308**(5728):1635-1638.

93. Berg R: **The indigenous gastrointestinal microflora**. *Trends Microbiol* 1996, **4**(11):430-435.

94. Ahmed S, Macfarlane GT, Fite A, McBain AJ, Gilbert P, Macfarlane S: **Mucosa-associated bacterial diversity in relation to human terminal ileum and colonic biopsy samples**. *Appl Environ Microbiol* 2007, **73**(22):7435-7442.

95. Hayashi H, Sakamoto M, Benno Y: **Phylogenetic analysis of the human gut microbiota using 16S rDNA clone libraries and strictly anaerobic culture-based methods**. *Microbiol Immunol* 2002, **46**(8):535-548.

96. Wang M, Ahrné S, Jeppsson B, Molin G: **Comparison of bacterial diversity along the human intestinal tract by direct cloning and sequencing of 16S rRNA genes**. *FEMS Microbiol Ecol* 2005, **54**(2):219-231.

97. Ben-Amor K, Heilig H, Smidt H, Vaughan EE, Abee T, de Vos WM: **Genetic diversity of viable, injured, and dead fecal bacteria assessed by fluorescence-activated cell sorting and 16S rRNA gene analysis**. *Appl Environ Microbiol* 2005, **71**(8):4679-4689.

98. Green GL, Brostoff J, Hudspith B, Michael M, Mylonaki M, Rayment N, Staines N, Sanderson J, Rampton DS, Bruce KD: **Molecular characterization of the bacteria adherent to human colorectal mucosa**. *J Appl Microbiol* 2006, **100**(3):460-469.

99. Vanhoutte T, Huys G, De Brandt E, Swings J: **Temporal stability analysis of the microbiota in human feces by denaturing gradient gel electrophoresis using universal and group-specific 16S rRNA gene primers.** *FEMS Microbiol Ecol* 2004, **48**:467-446.

100. Seksik P, Rigottier-Gois L, Gramet G, Sutren M, Pochart P, Marteau P, Jian R, Dore J: **Alterations of the dominant faecal bacterial groups in patients with Crohn's disease of the colon**. *Gut* 2003, **52**(2):237-242.

101. Favier CF, de Vos WM, Akkermans ADL: **Development of bacterial and bifidobacterial communities in feces of newborn babies**. *Anaerobe* 2003, **9**(5):219-229.

102. Ley RE, Peterson DA, Gordon JI: **Ecological and evolutionary forces shaping microbial diversity in the human intestine**. *Cell* 2006, **124**(4):837-848.

103. Rawls JF, Mahowald MA, Ley RE, Gordon JI: **Reciprocal gut microbiota transplants from zebrafish and mice to germ-free recipients reveal host habitat selection**. *Cell* 2006, **127**(2):423-433.

104. Gill SR, Pop M, DeBoy RT, Eckburg PB, Turnbaugh PJ, Samuel BS, Gordon JI, Relman DA, Fraser-Liggett CM, Nelson KE: **Metagenomic analysis of the human distal gut microbiome**. *Science* 2006, **312**(5778):1355-1359.

105. Chassard C, Scott KP, Marquet P, Martin JC, Del'homme C, Dapoigny M, Flint HJ, Bernalier-Donadille A: **Assessment of metabolic diversity within the intestinal microbiota from healthy humans using combined molecular and cultural approaches**. *FEMS Microbiol Ecol* 2008, **66**(3):496-504.

106. Shanahan F: **The host-microbe interface within the gut**. *Best Pract Res Clin Gastroenterol* 2002, **16**(6):915-931.

107. Clavel T, Haller D: **Bacteria- and host-derived mechanisms to control intestinal epithelial cell homeostasis: Implications for chronic inflammation**. *Inflamm Bowel Dis* 2007, **13**(9):1153-1164.

108. Hooper LV: **Bacterial contributions to mammalian gut development**. *Trends Microbiol* 2004, **12**(3):129-134.

109. Mueller C, Macpherson AJ: **Layers of mutualism with commensal bacteria protect us from intestinal inflammation**. *Gut* 2006, **55**(2):276-284.

110. Sartor RB: **Microbial influences in inflammatory bowel diseases**. *Gastroenterology* 2008, **134**(2):577-594.

111. Smythies LE, Sellers M, Clements RH, Mosteller-Barnum M, Meng G, Benjamin WH, Orenstein JM, Smith PD: **Human intestinal macrophages display profound inflammatory anergy despite avid phagocytic and bacteriocidal activity.** *J Clin Invest* 2005, **115**(1):66-75.

112. Backhed F, Ley RE, Sonnenburg JL, Peterson DA, Gordon JI: **Host-bacterial mutualism in the human intestine**. *Science* 2005, **307**(5717):1915-1920.

113. Xu J, Bjursell MK, Himrod J, Deng S, Carmichael LK, Chiang HC, Hooper L, Gordon JI: **A genomic view of the human-*Bacteroides thetaiotaomicron* symbiosis.** *Science* 2003, **299**(5615):2074-2076.

114. Sears C: **A dynamic partnership: celebrating our gut flora.** *Anaerobe* 2005, **11**(5):247-251.

115. An G, Wei B, Xia B, McDaniel JM, Ju T, Cummings RD, Braun J, Xia L: **Increased susceptibility to colitis and colorectal tumors in mice lacking core 3 derived O-glycans**. *J Exp Med* 2007, **204**(6):1417-1429.

116. Cario E: **Bacterial interactions with cells of the intestinal mucosa: Toll-like receptors and NOD2**. *Gut* 2005, **54**(8):1182-1193.

117. Xu Y, Tao X, Shen B, Horng T, Medzhitov R, Manley JL, Tong L: **Structural basis for signal transduction by the Toll/interleukin-1 receptor domains**. *Nature* 2000, **408**(6808):111-115.

118. Akira S, Takeda K: **Toll-like receptor signalling**. *Nat Rev Immunol* 2004, **4**(7):499-511.

119. Rosenstiel P, Jacobs G, Till A, Schreiber S: **NOD-like receptors: Ancient sentinels of the innate immune system**. *Cell Mol Life Sci* 2008, **65**(9):1361-1377.

120. Girardin SE, Boneca IG, Viala J, Chamaillard M, Labigne A, Thomas G, Philpott DJ, Sansonetti PJ: **Nod2 is a general sensor of peptidoglycan through muramyl dipeptide (MDP) detection**. *J Biol Chem* 2003, **278**(11):8869-8872.

121. Inohara N, Ogura Y, Fontalba A, Gutierrez O, Pons F, Crespo J, Fukase K, Inamura S, Kusumoto S, Hashimoto M, Foster SJ, Moran AP, Fernandez-Luna JL, Nunez G: **Host recognition of bacterial muramyl dipeptide mediated through NOD2. Implications for Crohn's disease.** *J Biol Chem* 2003, **278**(8):5509-5512.

122. Berrebi D, Maudinas R, Hugot JP, Chamaillard M, Chareyre F, De Lagausie P, Yang C, Desreumaux P, Giovannini M, Cezard JP, Zouali H, Emilie D, Peuchmaur M: **Card15 gene overexpression in mononuclear and epithelial cells of the inflamed Crohn's disease colon**. *Gut* 2003, **52**(6):840-846.

123. Rosenstiel P, Fantini M, Bräutigam K, Kühbacher T, Waetzig GH, Seegert D, Schreiber S: **TNF-[alpha] and IFN-[gamma] regulate the expression of the NOD2 (CARD15) gene in human intestinal epithelial cells**. *Gastroenterology* 2003, **124**(4):1001-1009.

124. Ogura Y, Lala S, Xin W, Smith E, Dowds TA, Chen FF, Zimmermann E, Tretiakova M, Cho JH, Hart J, Greenson JK, Keshav S, Nunez G: **Expression of NOD2 in Paneth cells: a possible link to Crohn's ileitis**. *Gut* 2003, **52**(11):1591-1597.

125. Bilsborough J, Viney JL: **Gastrointestinal dendritic cells play a role in immunity, tolerance, and disease**. *Gastroenterology* 2004, **127**(1):300-309.

126. Otte J-M, Rosenberg IM, Podolsky DK: **Intestinal myofibroblasts in innate immune responses of the intestine**. *Gastroenterology* 2003, **124**(7):1866-1878.

127. May G, Sutherland L, Meddings J: **Is small intestinal permeability really increased in relatives of patients with Crohn's disease?** *Gastroenterology* 1993, **104**(6):1627-1632.

128. Soderholm JD, Olaison G, Peterson KH, Franzen LE, Lindmark T, Wiren M, Tagesson C, Sjodahl R: **Augmented increase in tight junction permeability by luminal stimuli in the non-inflamed ileum of Crohn's disease**. *Gut* 2002, **50**(3):307-313.

129. Irvine EJ aMJ: **Increased intestinal permeability precedes the onset of Crohn's disease in a subject with familial risk.** *Gastroenterology* 2000, **119**(6):1740-1744.

130. Gassler N, Rohr C, Schneider A, Kartenbeck J, Bach A, Obermuller N, Otto HF, Autschbach F: **Inflammatory bowel disease is associated with changes of enterocytic junctions**. *Am J Physiol Gastrointest Liver Physiol* 2001, **281**(1):G216-228.

131. Buhner S, Buning C, Genschel J, Kling K, Herrmann D, Dignass A, Kuechler I, Krueger S, Schmidt HHJ, Lochs H: **Genetic basis for increased intestinal permeability in families with Crohn's disease: role of *CARD15* 3020insC mutation?** *Gut* 2006, **55**(3):342-347.

132. Heller F, Florian P, Bojarski C, Richter J, Christ M, Hillenbrand B, Mankertz J, Gitter AH, Bürgel N, Fromm M, Zeitz M, Fuss I, Strober W, Schulzke JD: **Interleukin-13 is the key effector Th2 cytokine in ulcerative colitis that affects epithelial tight junctions, apoptosis, and cell restitution**. *Gastroenterology* 2005, **129**(2):550-564.

133. Ma TY, Iwamoto GK, Hoa NT, Akotia V, Pedram A, Boivin MA, Said HM: **TNF-{alpha}-induced increase in intestinal epithelial tight junction permeability requires NF-{kappa}B activation**. *Am J Physiol Gastrointest Liver Physiol* 2004, **286**(3):G367-376.

134. Musch MW, Clarke LL, Mamah D, Gawenis LR, Zhang Z, Ellsworth W, Shalowitz D, Mittal N, Efthimiou P, Alnadjim Z, Hurst SD, Chang EB, Barrett TA: **T cell activation causes diarrhea by increasing intestinal permeability and inhibiting epithelial Na+/K+-ATPase**. *J Clin Invest* 2002, **110**(11):1739-1747.

135. Sun Y, Fihn BM, Sjovall H, Jodal M: **Enteric neurones modulate the colonic permeability response to luminal bile acids in rat colon in vivo**. *Gut* 2004, **53**(3):362-367.

136. Buisine MP, Desreumaux P, Leteurtre E, Copin MC, Colombel JF, Porchet N, Aubert JP: **Mucin gene expression in intestinal epithelial cells in Crohn's disease**. *Gut* 2001, **49**(4):544-551.

137. Smithson JE, Campbell A, Andrews JM, Milton JD, Pigott R, Jewell DP: **Altered expression of mucins throughout the colon in ulcerative colitis**. *Gut* 1997, **40**(2):234-240.

138. Itoh H, Beck PL, Inoue N, Xavier R, Podolsky DK: **A paradoxical reduction in susceptibility to colonic injury upon targeted transgenic ablation of goblet cells.** *J Clin Invest* 1999 **104**(11):1539-1547.

139. Van der Sluis M, De Koning BAE, De Bruijn ACJM, Velcich A, Meijerink JPP, Van Goudoever JB, Buller HA, Dekker J, Van Seuningen I, Renes IB, Einerhand AWC: **Muc2-deficient mice spontaneously develop colitis, indicating that MUC2 is critical for colonic protection**. *Gastroenterology* 2006, **131**(1):117-129.

140. Korzenik JR: **Is Crohn's disease due to defective immunity?** *Gut* 2007, **56**(1):2-5.

141. Duchmann R, Kaiser I, Hermann E, Mayet W, Ewe K, Meyer zum Büschenfelde KH: **Tolerance exists towards resident intestinal flora but is broken in active inflammatory bowel disease (IBD)**. *Clin Exp Immunol* 1995 **102**(3):448-455.

142. Cario E, Podolsky DK: **Differential alteration in intestinal epithelial cell expression of toll-like receptor 3 (TLR3) and TLR4 in inflammatory bowel disease**. *Infect Immun* 2000, **68**(12):7010-7017

143. Hart AL, Al-Hassi HO, Rigby RJ, Bell SJ, Emmanuel AV, Knight SC, Kamm MA, Stagg AJ: **Characteristics of intestinal dendritic cells in inflammatory bowel diseases**. *Gastroenterology* 2005, **129**(1):50-65.

144. Wehkamp J, Salzman NH, Porter E, Nuding S, Weichenthal M, Petras RE, Shen B, Schaeffeler E, Schwab M, Linzmeier R, Feathers RW, Chu H, Lima H, Fellermann K, Ganz T, Stange EF, Bevins CL: **Reduced Paneth cell (alpha)-defensins in ileal Crohn's disease**. *Proc Natl Acad Sci USA* 2005, **102**(50):18129-18134.

145. Kobayashi KS, Chamaillard M, Ogura Y, Henegariu O, Inohara N, Nunez G, Flavell RA: **Nod2-dependent regulation of innate and adaptive immunity in the intestinal tract**. *Science* 2005, **307**(5710):731-734.

146. Wehkamp J, Harder J, Weichenthal M, Schwab M, Schaffeler E, Schlee M, Herrlinger KR, Stallmach A, Noack F, Fritz P, Schroder JM, Bevins CL, Fellermann K, Stange EF: **NOD2 (CARD15) mutations in Crohn's disease are associated with diminished mucosal {alpha}-defensin expression**. *Gut* 2004, **53**(11):1658-1664.

147. Wehkamp J, Wang G, Kubler I, Nuding S, Gregorieff A, Schnabel A, Kays RJ, Fellermann K, Burk O, Schwab M, Clevers H, Bevins CL, Stange EF: **The Paneth cell {alpha}-defensin deficiency of ileal Crohn's disease is linked to Wnt/Tcf-4**. *J Immunol* 2007, **179**(5):3109-3118.

148. Ayabe T, Satchell D, Wilson C, Parks W, Selsted M, Ouellette A: **Secretion of microbicidal alpha-defensins by intestinal Paneth cells in response to bacteria**. *Nat Immunol* 2000, **1**(2):113-118.

149. Salzman NH, Underwood MA, Bevins CL: **Paneth cells, defensins, and the commensal microbiota: A hypothesis on intimate interplay at the intestinal mucosa**. *Semin Immunol* 2007, **19**(2):70-83.

150. Wehkamp J, Fellermann K, Herrlinger KR, Baxmann S, Schmidt K, Schwind B, Duchrow M, Wohlschläger C, Feller A, Stange EF: **Human beta-defensin 2 but not beta-defensin 1 is expressed preferentially in colonic mucosa of inflammatory bowel disease.** *Eur J Gastroenterol Hepatol* 2002 **14**(7):745-752.

151. Wehkamp J, Harder J, Weichenthal M, Mueller O, Herrlinger KR, Fellermann K, Schroeder JM, Stange EF: **Inducible and constitutive beta-defensins are differentially expressed in Crohn's disease and ulcerative colitis.** *Inflamm Bowel Dis* 2003 **9**(4):215-223.

152. Fellermann K, Stange DE, Schaeffeler E, Schmalzl H, Wehkamp J, Bevins CL, Reinisch W, Teml A, Schwab M, Lichter P, Radlwimmer B, Stange EF: **A chromosome 8 gene-cluster polymorphism with low human beta-defensin 2 gene copy number predisposes to Crohn disease of the colon.** *Am J Hum Genet* 2006, **79**(3):439-448.

153. Bonen DK, Ogura Y, Nicolae DL, Inohara N, Saab L, Tanabe T, Chen FF, Foster SJ, Duerr RH, Brant SR, Cho JH, Nuñez G: **Crohn's disease-associated NOD2 variants share a signaling defect in response to lipopolysaccharide and peptidoglycan**. *Gastroenterology* 2003, **124**(1):140-146.

154. Zelinkova Z, van Beelen AJ, de Kort F, Moerland PD, van Themaat EV, te Velde AA, van Deventer SJ, de Jong EC, Hommes DW: **Muramyl dipeptide-induced differential gene expression in NOD2 mutant and wild-type crohn's disease patient-derived dendritic cells**. *Inflamm Bowel Dis* 2008, **14**(2):186-194.

155. Chamaillard M, Philpott D, Girardin SE, Zouali H, Lesage S, Chareyre F, Bui TH, Giovannini M, Zaehringer U, Penard-Lacronique V, Sansonetti PJ, Hugot J-P, Thomas G: **Gene-environment interaction modulated by allelic heterogeneity in inflammatory diseases**. *Proc Natl Acad Sci USA* 2003, **100**(6):3455-3460.

156. Kramer M, Netea MG, de Jong DJ, Kullberg BJ, Adema GJ: **Impaired dendritic cell function in Crohn's disease patients with NOD2 3020insC mutation**. *J Leukoc Biol* 2006, **79**(4):860-866.

157. Netea MG, Ferwerda G, de Jong DJ, Jansen T, Jacobs L, Kramer M, Naber THJ, Drenth JPH, Girardin SE, Jan Kullberg B, Adema GJ, Van der Meer JWM: **Nucleotide-binding oligomerization domain-2 modulates specific TLR pathways for the induction of cytokine release**. *J Immunol* 2005, **174**(10):6518-6523.

158. van Heel DA, Ghosh S, Hunt KA, Mathew CG, Forbes A, Jewell DP, Playford RJ: **Synergy between TLR9 and NOD2 innate immune responses is lost in genetic Crohn's disease**. *Gut* 2005, **54**(11):1553-1557.

159. Maeda S, Hsu L-C, Liu H, Bankston LA, Iimura M, Kagnoff MF, Eckmann L, Karin M: ***Nod2* mutation in Crohn's disease potentiates NF-{kappa}B activity and IL-1{beta} processing**. *Science* 2005, **307**(5710):734-738.

160. Watanabe T, Kitani A, Murray PJ, Strober W: **NOD2 is a negative regulator of Toll-like receptor 2-mediated T helper type 1 responses**. *Nat Immunol* 2004, **5**(8):800-808.

161. Strober W, Fuss I, Mannon P: **The fundamental basis of inflammatory bowel disease**. *J Clin Invest* 2007, **117**(3):514-521.

162. Watanabe T, Kitani A, Murray Peter J, Wakatsuki Y, Fuss IJ, Strober W: **Nucleotide Binding Oligomerization Domain 2 deficiency leads to dysregulated TLR2 signaling and induction of antigen-specific colitis**. *Immunity* 2006, **25**(3):473-485.

163. Cruickshank S, McVay L, Baumgart D, Felsburg P, Carding S: **Colonic epithelial cell mediated suppression of CD4 T cell activation**. *Gut* 2004, **53**(5):678-684.

164. Nakazawa A, Dotan I, Brimnes J, Allez M, Shao L, Tsushima F, Azuma M, Mayer L: **The expression and function of costimulatory molecules B7H and B7-H1 on colonic epithelial cells**. *Gastroenterology* 2004, **126**(5):1347-1357.

165. Gewirtz AT: **TLRs in the Gut. III. Immune responses to flagellin in Crohn's disease: good, bad, or irrelevant?** *Am J Physiol Gastrointest Liver Physiol* 2007, **292**(3):G706-710.

166. Didierlaurent A, Ferrero I, Otten LA, Dubois B, Reinhardt M, Carlsen H, Blomhoff R, Akira S, Kraehenbuhl J-P, Sirard J-C: **Flagellin promotes myeloid differentiatiofactor 88-dependent development of Th2-type response**. *J Immunol* 2004, **172**(11):6922-6930.

167. McSorley SJ, Ehst BD, Yu Y, Gewirtz AT: **Bacterial flagellin is an effective adjuvant for CD4+ T cells *in vivo***. *J Immunol* 2002, **169**(7):3914-3919.

168. Ina K, Itoh J, Fukushima K, Kusugami K, Yamaguchi T, Kyokane K, Imada A, Binion DG, Musso A, West GA, Dobrea GM, McCormick TS, Lapetina EG, Levine AD, Ottaway CA, Fiocchi C: **Resistance of Crohn's disease T cells to multiple apoptotic signals is associated with a Bcl-2/Bax mucosal imbalance**. *J Immunol* 1999, **163**(2):1081-1090.

169. Van den Brande JMH, Braat H, van den Brink GR, Versteeg HH, Bauer CA, Hoedemaeker I, van Montfrans C, Hommes DW, Peppelenbosch MP, van Deventer SJH: **Infliximab but not etanercept induces apoptosis in lamina propria T-lymphocytes from patients with Crohn's disease**. *Gastroenterology* 2003, **124**(7):1774-1785.

170. Fuss IJ, Heller F, Boirivant M, Leon F, Yoshida M, Fichtner-Feigl S, Yang Z, Exley M, Kitani A, Blumberg RS, Mannon P, Strober W: **Nonclassical CD1d-restricted NK T cells that produce IL-13 characterize an atypical Th2 response in ulcerative colitis.** *J Clin Invest* 2004, **113**(10):1490-1497.

171. Fuss IJ, Neurath M, Boirivant M, Klein JS, de la Motte C, Strong SA, Fiocchi C, Strober W: **Disparate CD4+ lamina propria (LP) lymphokine secretion profiles in inflammatory bowel disease. Crohn's disease LP cells manifest increased secretion of IFN-gamma, whereas ulcerative colitis LP cells manifest increased secretion of IL-5**. *J Immunol* 1996, **157**(3):1261-1270.

172. Langrish CL, Chen Y, Blumenschein WM, Mattson J, Basham B, Sedgwick JD, McClanahan T, Kastelein RA, Cua DJ: **IL-23 drives a pathogenic T cell population that induces autoimmune inflammation**. *J Exp Med* 2005, **201**(2):233-240.

173. Neurath MF, Weigmann B, Finotto S, Glickman J, Nieuwenhuis E, Iijima H, Mizoguchi A, Mizoguchi E, Mudter J, Galle PR, Bhan A, Autschbach F, Sullivan BM, Szabo SJ, Glimcher LH, Blumberg RS: **The transcription factor T-bet regulates mucosal T cell activation in experimental colitis and Crohn's disease**. *J Exp Med* 2002, **195**(9):1129-1143.

174. Plevy SE, Landers CJ, Prehn J, Carramanzana NM, Deem RL, Shealy D, Targan SR: **A role for TNF-alpha and mucosal T helper-1 cytokines in the pathogenesis of Crohn's disease**. *J Immunol* 1997, **159**(12):6276-6282.

175. Brand S, Beigel F, Olszak T, Zitzmann K, Eichhorst ST, Otte J-M, Diepolder H, Marquardt A, Jagla W, Popp A, Leclair S, Herrmann K, Seiderer J, Ochsenkuhn T, Goke B, Auernhammer CJ, Dambacher J: **IL-22 is increased in active Crohn's disease and promotes proinflammatory gene expression and intestinal epithelial cell migration**. *Am J Physiol Gastrointest Liver Physiol* 2006, **290**(4):G827-838.

176. Seiderer J, Elben I, Diegelmann J, Glas J, Stallhofer J, Tillack C, Pfennig S, Jürgens M, Schmechel S, Konrad A, Göke B, Ochsenkühn T, Müller-Myhsok B, Lohse P, Brand S: **Role of the novel Th17 cytokine IL-17F in inflammatory bowel disease (IBD): Upregulated colonic IL-17F expression in active Crohn's disease and analysis of the *IL17F* p.His161Arg polymorphism in IBD**. *Inflamm Bowel Dis* 2008, **14**(4):437-445.

177. Yen D, Cheung J, Scheerens H, Poulet F, McClanahan T, McKenzie B, Kleinschek M, Owyang A, Mattson J, Blumenschein W, Murphy E, Sathe M, Cua D, Kastelein R, Rennick D: **IL-23 is essential for T cell–mediated colitis and promotes inflammation via IL-17 and IL-6**. *J Clin Invest* 2006, **116**(5):1310–1316.

178. Zhang Z, Hinrichs DJ, Lu H, Chen H, Zhong W, Kolls JK: **After interleukin-12p40, are interleukin-23 and interleukin-17 the next therapeutic targets for inflammatory bowel disease?** *Int Immunopharmacol* 2007, **7**(4):409-416.

179. Schmechel S, Konrad A, Diegelmann J, Glas J, Wetzke M, Paschos E, Lohse P, Göke B, Brand S: **Linking genetic susceptibility to Crohn's disease with Th17 cell function: IL-22 serum levels are increased in Crohn's disease and correlate with disease activity and *IL23R* genotype status**. *Inflamm Bowel Dis* 2008, **14**(2):204-212.

180. Dicksved J, Halfvarson J, Rosenquist M, Jarnerot G, Tysk C, Apajalahti J, Engstrand L, Jansson JK: **Molecular analysis of the gut microbiota of identical twins with Crohn's disease**. *ISME J* 2008.

181. Harper PH, Lee EC, Kettlewell MG, Bennett MK, Jewell DP: **Role of the faecal stream in the maintenance of Crohn's colitis**. *Gut* 1985, **26**(3):279-284.

182. Rutgeerts P, Goboes K, Peeters M, Hiele M, Penninckx F, Aerts R, Kerremans R, Vantrappen G: **Effect of faecal stream diversion on recurrence of Crohn's disease in the neoterminal ileum.** *Lancet* 1991, **338**(8770):771-774.

183. Rutgeerts P, Hiele M, Geboes K, Peeters M, Penninckx F, Aerts R, Kerremans R: **Controlled trial of metronidazole treatment for prevention of crohn's recurrence after ileal resection**. *Gastroenterology* 1995, **108**(6):1617-1621.

184. Sartor RB, Rath HC, Lichtman SN, Van Tol EAF: **Animal models of intestinal and joint inflammation**. *Baillieres Clin Rheumatol* 1996, **10**(1):55-76.

185. Abrams R: **Open-label, uncontrolled trial of bowel sterilization and repopulation with normal bowel flora for treatment of inflammatory bowel disease** *Curr Ther Res* 1997, **58**(12):1001-1012.

186. Quinton JF, Sendid B, Reumaux D, Duthilleul P, Cortot A, Grandbastien B, Charrier G, Targan SR, Colombel JF, Poulain D: **Anti-*Saccharomyces cerevisiae* mannan antibodies combined with antineutrophil cytoplasmic autoantibodies in inflammatory bowel disease: prevalence and diagnostic role**. *Gut* 1998, **42**(6):788-791.

187. Schultsz C, van den Berg FM, ten Kate FW, Tytgat GNJ, Dankert J: **The intestinal mucus layer from patients with inflammatory bowel disease harbors high numbers of bacteria compared with controls**. *Gastroenterology* 1999, **117**(5):1089-1097.

188. Prindiville T, Sheikh R, Cohen S, Tang Y, Cantrell M, Silva JJ: ***Bacteroides fragilis* enterotoxin gene sequences in patients with inflammatory bowel disease.** *Emerg Infect Dis* 2000, **6**(2):171-174.

189. Sutton CL, Kim J, Yamane A, Dalwadi H, Wei B, Landers C, Targan SR, Braun J: **Identification of a novel bacterial sequence associated with Crohn's disease**. *Gastroenterology* 2000, **119**(1):23-31.

190. Wei B, Huang T, Dalwadi H, Sutton CL, Bruckner D, Braun J: ***Pseudomonas fluorescens* encodes the Crohn's disease-associated I2 sequence and T-cell superantigen**. *Infect Immun* 2002, **70**(12):6567-6575.

191. Landers CJ, Cohavy O, Misra R, Yang H, Lin Y-C, Braun J, Targan SR: **Selected loss of tolerance evidenced by Crohn's disease-associated immune responses to auto- and microbial antigens**. *Gastroenterology* 2002, **123**(3):689-699.

192. Mow WS, Vasiliauskas EA, Lin Y-C, Fleshner PR, Papadakis KA, Taylor KD, Landers CJ, Abreu-Martin MT, Rotter JI, Yang H, Targan SR: **Association of antibody responses to microbial antigens and complications of small bowel Crohn's disease**. *Gastroenterology* 2004, **126**(2):414-424.

193. Swidsinski A, Ladhoff A, Pernthaler A, Swidsinski S, Loening-Baucke V, Ortner M, Weber J, Hoffmann U, Schreiber S, Dietel M, Lochs H: **Mucosal flora in inflammatory bowel disease**. *Gastroenterology* 2002, **122**(1):44-54.

194. Kleessen B, Kroesen A, Buhr H, Blaut M: **Mucosal and invading bacteria in patients with inflammatory bowel disease compared with controls.** *Scand J Gastroenterol* 2002, **37**(9):1034-1041.

195. Seksik P, Sokol H, Lepage P, Vasquez N, Manichanh C, Mangin I, Pochart P, Doré J, Marteau P: **Review article: the role of bacteria in onset and perpetuation of inflammatory bowel disease**. *Aliment Pharmacol Ther* 2006, **24**(s3):11-18.

196. Tamboli CP, Neut C, Desreumaux P, Colombel JF: **Dysbiosis in inflammatory bowel disease**. *Gut* 2004, **53**(1):1-4.

197. Chen W, Li D, Paulus B, Wilson I, Chadwick VS: **Detection of *Listeria monocytogenes* by polymerase chain reaction in intestinal mucosal biopsies from patients with inflammatory bowel disease and controls**. *J Gastroenterol Hepatol* 2000, **15**(10):1145-1150.

198. Hermon-Taylor J, Bull T, Sheridan J, Cheng J, Stellakis M, Sumar N: **Causation of Crohn's disease by *Mycobacterium avium* subspecies paratuberculosis.** *Can J Gastroenterol* 2000, **14**(6):521-539.

199. Lamps LW, Madhusudhan KT, Havens JM, Greenson JK, Bronner MP, Chiles MC, Dean PJ, Scott MA: **Pathogenic *Yersinia* DNA is detected in bowel and mesenteric lymph nodes from patients with Crohn's disease**. *Am J Surg Pathol* 2003 **27**(2):220-227.

200. Liu Y, van Kruiningen H, West A, Cartun R, Cortot A, Colombel J: **Immunocytochemical evidence of *Listeria*, *Escherichia coli*, and *Streptococcus* antigens in Crohn's disease.** *Gastroenterology* 1995 **108**(5):1396-1404.

201. O'Mahony S, Anderson N, Nuki G, Ferguson A: **Systemic and mucosal antibodies to *Klebsiella* in patients with ankylosing spondylitis and Crohn's disease**. *Ann Rheum Dis* 1992, **51**(12):1296-1300.

202. Tabaqchali S, O'Donoghue DP, Bettelheim KA: ***Escherichia coli* antibodies in patients with inflammatory bowel disease**. *Gut* 1978, **19**(2):108-113.

203. Bibiloni R, Mangold M, Madsen KL, Fedorak RN, Tannock GW: **The bacteriology of biopsies differs between newly diagnosed, untreated, Crohn's disease and ulcerative colitis patients**. *J Med Microbiol* 2006, **55**(8):1141-1149.

204. Vasquez N, Mangin I, Lepage P, Seksik P, Duong J-P, Blum S, Schiffrin E, Suau A, Allez M, Vernier G, Tréton X, Doré J, Marteau P, Pochart P: **Patchy distribution of mucosal lesions in ileal Crohn's disease is not linked to differences in the dominant mucosa-associated bacteria: A study using fluorescence in situ hybridization and temporal temperature gradient gel electrophoresis**. *Inflamm Bowel Dis* 2007, **13**(6):684-692.

205. Martinez-Medina M, Aldeguer X, Gonzalez-Huix F, Acero D, Garcia-Gil LJ: **Abnormal microbiota composition in the ileocolonic mucosa of Crohn's**

disease patients as revealed by polymerase chain reaction-denaturing gradient gel electrophoresis. *Inflamm Bowel Dis* 2006, **12**(12):1136-1145.

206. Seksik P, Lepage P, de la Cochetiere M-F, Bourreille A, Sutren M, Galmiche J-P, Dore J, Marteau P: **Search for localized dysbiosis in Crohn's disease ulcerations by temporal temperature gradient gel electrophoresis of 16S rRNA**. *J Clin Microbiol* 2005, **43**(9):4654-4658.

207. Baumgart M, Dogan B, Rishniw M, Weitzman G, Bosworth B, Yantiss R, Orsi RH, Wiedmann M, McDonough P, Kim SG, Berg D, Schukken Y, Scherl E, Simpson KW: **Culture independent analysis of ileal mucosa reveals a selective increase in invasive *Escherichia coli* of novel phylogeny relative to depletion of Clostridiales in Crohn's disease involving the ileum**. *ISME J* 2007, **1**(5):403-418.

208. Lepage P, Seksik P, Sutren M, de la Cochetière M-F, Jian R, Marteau P, Doré J: **Biodiversity of the mucosa-associated microbiota is stable along the distal digestive tract in healthy individuals and patients with IBD**. *Inflamm Bowel Dis* 2005, **11**(5):473-480.

209. Ott SJ, Musfeldt M, Wenderoth DF, Hampe J, Brant O, Folsch UR, Timmis KN, Schreiber S: **Reduction in diversity of the colonic mucosa associated bacterial microflora in patients with active inflammatory bowel disease**. *Gut* 2004, **53**(5):685-693.

210. Scanlan PD, Shanahan F, O'Mahony C, Marchesi JR: **Culture-independent analyses of temporal variation of the dominant fecal microbiota and targeted bacterial subgroups in Crohn's disease**. *J Clin Microbiol* 2006, **44**(11):3980-3988.

211. Andoh A, Tsujikawa T, Sasaki M, Mitsuyama K, Suzuki Y, Matsui T, Matsumoto T, Benno Y, Fujiyama Y: **Fecal microbiota profile of crohn's disease determined by terminal restriction fragment length polymorphism (t-rflp) analysis. In press.** *Aliment Pharmacol Ther* 2008, **29**(1):75-82.

212. Gophna U, Sommerfeld K, Gophna S, Doolittle WF, Veldhuyzen van Zanten SJO: **Differences between tissue-associated intestinal microfloras of patients with Crohn's disease and ulcerative colitis**. *J Clin Microbiol* 2006, **44**(11):4136-4141.

213. Manichanh C, Rigottier-Gois L, Bonnaud E, Gloux K, Pelletier E, Frangeul L, Nalin R, Jarrin C, Chardon P, Marteau P, Roca J, Dore J: **Reduced diversity of faecal microbiota in Crohn's disease revealed by a metagenomic approach**. *Gut* 2006, **55**(2):205-211.

214. Prindiville T, Cantrell M, Wilson KH: **Ribosomal DNA sequence analysis of mucosa-associated bacteria in Crohn's disease**. *Inflamm Bowel Dis* 2004, **10**(6):824-833.

215. Swidsinski A, Loening-Baucke V, Vaneechoutte M, Doerffel Y: **Active Crohn's disease and ulcerative colitis can be specifically diagnosed and monitored based on the biostructure of the fecal flora**. *Inflamm Bowel Dis* 2008, **14**(2):147-161.

216. Swidsinski A, Weber J, Loening-Baucke V, Hale LP, Lochs H: **Spatial organization and composition of the mucosal flora in patients with inflammatory bowel disease**. *J Clin Microbiol* 2005, **43**(7):3380-3389.

217. Takaishi H, Matsuki T, Nakazawa A, Takada T, Kado S, Asahara T, Kamada N, Sakuraba A, Yajima T, Higuchi H, Inoue N, Ogata H, Iwao Y, Nomoto K, Tanaka R, Hibi T: **Imbalance in intestinal microflora constitution could be involved in the pathogenesis of inflammatory bowel disease**. *Int J Med Microbiol* 2008, **298**(5-6):463-472.

218. Frank DN, St. Amand AL, Feldman RA, Boedeker EC, Harpaz N, Pace NR: **Molecular-phylogenetic characterization of microbial community imbalances in human inflammatory bowel diseases**. *Proc Natl Acad Sci USA* 2007, **104**(34):13780-13785.

219. Willing B, Halfvarson J, Dicksved J, Rosenquist M, Järnerot G, Engstrand L, Tysk C, Jansson JK: **Twin studies reveal specific imbalances in the mucosa-associated microbiota of patients with ileal Crohn's disease**. *Inflamm Bowel Dis* 2009, **15**(5):653-660.

220. Conte MP, Schippa S, Zamboni I, Penta M, Chiarini F, Seganti L, Osborn J, Falconieri P, Borrelli O, Cucchiara S: **Gut-associated bacterial microbiota in paediatric patients with inflammatory bowel disease**. *Gut* 2006, **55**(12):1760-1767.

221. Favier C, Neut C, Mizon C, Cortot A, Colombel J, Mizon J: **Fecal beta-D-galactosidase production and *Bifidobacteria* are decreased in Crohn's disease.** *Dig Dis Sci* 1997, **42**(4):817-822.

222. Kuehbacher T, Rehman A, Lepage P, Hellmig S, Folsch UR, Schreiber S, Ott SJ: **Intestinal TM7 bacterial phylogenies in active inflammatory bowel disease**. *J Med Microbiol* 2008, **57**(12):1569-1576.

223. Ott SJ, Kühbacher T, Musfeldt M, Rosenstiel P, Hellmig S, Rehman A, Drews O, Weichert W, Timmis KN, Schreiber S: **Fungi and inflammatory bowel diseases: Alterations of composition and diversity**. *Scand J Gastroenterol* 2008, **43**(7):831 - 841.

224. De Hertogh G, Aerssens J, Geboes K, Geboes K: **Evidence for the involvement of infectious agents in the pathogenesis of Crohn's disease**. *World J Gastroenterol* 2008 **14**(6):845-852.

225. Chiba M, Kono M, Hoshina S, Komatsu M, Kitagawa Y, Iizuka M, Watanabe S: **Presence of bacterial 16S ribosomal RNA gene segments in human intestinal lymph follicles.** *Scand J Gastroenterol* 2000 **35**(8):824-831.

226. Walmsley RS, Anthony A, Sim R, Pounder RE, Wakefield AJ: **Absence of *Escherichia coli*, *Listeria monocytogenes*, and *Klebsiella pneumoniae* antigens within inflammatory bowel disease tissues**. *J Clin Pathol* 1998, **51**(9):657-661.

227. Williams G, Khan A, Schweiger F: ***Listeria meningitis* complicating infliximab treatment for Crohn's disease.** *Can J Infect Dis Med Microbiol* 2005 **16**(5):289-292.

228. Chiodini R, Van Kruiningen H, Thayer W, Merkal R, Coutu J: **Possible role of mycobacteria in inflammatory bowel disease. I. An unclassified *Mycobacterium* species isolated from patients with Crohn's disease.** *Dig Dis Sci* 1984, **29**(12):1073-1079.

229. Abubakar I, Myhill D, Aliyu SH, Hunter PR: **Detection of *Mycobacterium avium* subspecies *paratuberculosis* from patients with Crohn's disease using nucleic acid-based techniques: A systematic review and meta-analysis**. *Inflamm Bowel Dis* 2008, **14**(3):401-410.

230. Bentley R, Keenan J, Gearry R, Kennedy M, Barclay M, Roberts R: **Incidence of *Mycobacterium avium* subspecies *paratuberculosis* in a population-based cohort of patients with Crohn's disease and control subjects.** *Am J Gastroenterol* 2008, **103**(5):1168-1172.

231. Bull TJ, McMinn EJ, Sidi-Boumedine K, Skull A, Durkin D, Neild P, Rhodes G, Pickup R, Hermon-Taylor J: **Detection and verification of *Mycobacterium avium* subsp. *paratuberculosis* in fresh ileocolonic mucosal biopsy specimens from**

individuals with and without Crohn's disease. *J Clin Microbiol* 2003, **41**(7):2915-2923.

232. Naser SA, Ghobrial G, Romero C, Valentine JF: **Culture of *Mycobacterium avium* subspecies *paratuberculosis* from the blood of patients with Crohn's disease**. *Lancet* 2004, **364**(9439):1039-1044.

233. Ryan P, Bennett MW, Aarons S, Lee G, Collins JK, O'Sullivan GC, O'Connell J, Shanahan F: **PCR detection of *Mycobacterium paratuberculosis* in Crohn's disease granulomas isolated by laser capture microdissection**. *Gut* 2002, **51**(5):665-670.

234. Scanu AM, Bull TJ, Cannas S, Sanderson JD, Sechi LA, Dettori G, Zanetti S, Hermon-Taylor J: ***Mycobacterium avium* subspecies *paratuberculosis* infection in cases of irritable bowel syndrome and comparison with Crohn's disease and Johne's disease: Common neural and immune pathogenicities**. *J Clin Microbiol* 2007, **45**(12):3883-3890.

235. Sechi LA, Scanu AM, Molicotti P, Cannas S, Mura M, Dettori G, Fadda G, Zanetti S: **Detection and isolation of *Mycobacterium avium* subspecies *paratuberculosis* from intestinal mucosal biopsies of patients with and without Crohn's disease in Sardinia**. *Am J Gastroenterol* 2005, **100**(7):1529-1536.

236. Selby W, Pavli P, Crotty B, Florin T, Radford-Smith G, Gibson P, Mitchell B, Connell W, Read R, Merrett M, Ee H, Hetzel D: **Two-year combination antibiotic therapy with clarithromycin, rifabutin, and clofazimine for Crohn's disease**. *Gastroenterology* 2007, **132**(7):2313-2319.

237. Feller M, Huwiler K, Stephan R, Altpeter E, Shang A, Furrer H, Pfyffer GE, Jemmi T, Baumgartner A, Egger M: ***Mycobacterium avium* subspecies *paratuberculosis* and Crohn's disease: a systematic review and meta-analysis**. *Lancet Infect Dis* 2007, **7**(9):607-613.

238. Bohr URM, Glasbrenner B, Primus A, Zagoura A, Wex T, Malfertheiner P: **Identification of enterohepatic *Helicobacter* species in patients suffering from inflammatory bowel disease**. *J Clin Microbiol* 2004, **42**(6):2766-2768.

239. Oliveira AG, Rocha GA, Rocha AMC, Sanna MdGP, Moura SB, Dani R, Marinho FP, Moreira LS, Ferrari MdLA, Castro LPF, Queiroz DMM: **Isolation of *Helicobacter pylori* from the intestinal mucosa of patients with Crohn's disease**. *Helicobacter* 2006, **11**(1):2-9.

240. Jergens AE, Wilson-Welder JH, Dorn A, Henderson A, Liu Z, Evans RB, Hostetter J, Wannemuehler MJ: ***Helicobacter bilis* triggers persistent immune reactivity to antigens derived from the commensal bacteria in gnotobiotic C3H/HeN mice**. *Gut* 2007, **56**(7):934-940.

241. Jawhara S, Thuru X, Standaert-Vitse A, Jouault T, Mordon S, Sendid B, Desreumaux P, Poulain D: **Colonization of mice by *Candida albicans* is promoted by chemically induced colitis and augments inflammatory responses through Galectin-3**. *J Infect Dis* 2008, **197**(7):972-980.

242. Standaert-Vitse A, Jouault T, Vandewalle P, Mille C, Seddik M, Sendid B, Mallet J, Colombel J, Poulain D: ***Candida albicans* is an immunogen for anti-*Saccharomyces cerevisiae* antibody markers of Crohn's disease.** *Gastroenterology* 2006 **130**(6):1764-1775.

243. Cruyssen BV, Peeters H, Hoffman EA, Laukens D, Coucke P, Marichal D, Cuvelier C, Remaut E, Veys EM, Mielants H, De Vos M, De Keyser F: ***CARD15* polymorphisms are associated with anti-*Saccharomyces cerevisiae* antibodies in caucasian Crohn's disease patients.** *Clin Exp Immunol* 2005 **140**(2):354-359. .

244. Dassopoulos T, Frangakis C, Cruz-Correa M, Talor MV, Burek CL, Datta L, Nouvet F, Bayless TM, Brant SR: **Antibodies to *Saccharomyces cerevisiae* in Crohn's disease: Higher titers are associated with a greater frequency of mutant *NOD2/CARD15* alleles and with a higher probability of complicated disease**. *Inflamm Bowel Dis* 2007, **13**(2):143-151.

245. Papp M, Altorjay I, Norman GL, Shums Z, Palatka K, Vitalis Z, Foldi I, Lakos G, Tumpek J, Udvardy ML, Harsfalvi J, Fischer S, Lakatos L, Kovacs A, Bene L, Molnar T, Tulassay Z, Miheller P, Veres G, Papp J: **Seroreactivity to microbial components in Crohn's disease is associated with ileal involvement, noninflammatory disease behavior and *NOD2/CARD15* genotype, but not with risk for surgery in a Hungarian cohort of IBD patients**. *Inflamm Bowel Dis* 2007, **13**(8):984-992.

246. Wolters F, Russel M, Sijbrandij J, Schouten L, Odes S, Riis L, Munkholm P, Langholz E, Bodini P, O'Morain C, Katsanos K, Tsianos E, Vermeire S, Van Zeijl G, Limonard C, Hoie O, Vatn M, Moum B, Stockbrügger R: **Disease outcome of inflammatory bowel disease patients: general outline of a Europe-wide population-based 10-year clinical follow-up study.** *Scand J Gastroenterol Suppl* 2006 **243**:46-54.

247. Seibold F, Konrad A, Flogerzi B, Seibold-Schmid B, Arni S, Jüliger S, Kun J: **Genetic variants of the mannan-binding lectin are associated with immune reactivity to mannans in Crohn's disease.** *Gastroenterology* 2004 **127**(4):1076-1084.

248. Arnott IDR, Landers CJ, Nimmo EJ, Drummond HE, Smith BKR, Targan SR, Satsangi J: **Sero-reactivity to microbial components in Crohn's disease is associated with disease severity and progression, but not NOD2/CARD15 genotype**. *Am J Gastroenterol* 2004, **99**(12):2376-2384.

249. Hadrich I, Vandewalle P, Cheikhrouhou F, Makni F, Krichen MS, Sendid B, Standaert-Vitse A, Ayadi A, Poulain D: **Ethnic and socio-cultural specificities in Tunisia have no impact on the prevalence of anti-*Saccharomyces cerevisiae* antibodies in Crohn's disease patients, their relatives or associated clinical factors**. *Scand J Gastroenterol* 2007, **42**:717-725.

250. Joossens S, Pierik M, Rector A, Vermeire S, Ranst MV, Rutgeerts P, Bossuyt X: **Mannan binding lectin (MBL) gene polymorphisms are not associated with anti-*Saccharomyces cerevisiae* (ASCA) in patients with Crohn's disease**. *Gut* 2006, **55**(5):746-.

251. Gorbach SL, Nahas L, Plaut AG, Weinstein L, Patterson JF, Levitan R: **Studies of intestinal microflora. V. Fecal microbial ecology in ulcerative colitis and regional enteritis: Relationship to severity of disease and chemotherapy**. *Gastroenterology* 1968, **54**(4):575.

252. Keighley MR, Arabi Y, Dimock F, Burdon DW, Allan RN, Alexander-Williams J: **Influence of inflammatory bowel disease on intestinal microflora**. *Gut* 1978, **19**(12):1099-1104.

253. Neut C, Colombel JF, Guillemot F, Cortot A, Gower P, Quandalle P, Ribet M, Romond C, Paris JC: **Impaired bacterial flora in human excluded colon**. *Gut* 1989, **30**(8):1094-1098.

254. Fujita H, Eishi Y, Ishige I, Saitoh K, Takizawa T, Arima T, Koike M: **Quantitative analysis of bacterial DNA from *Mycobacteria* spp., *Bacteroides vulgatus*, and *Escherichia coli* in tissue samples from patients with inflammatory bowel diseases**. *J Gastroenterol* 2002, **37**(7):509-516.

255. Mangin I, Bonnet R, Seksik P, Rigottier-Gois L, Sutren M, Bouhnik Y, Neut C, Collins MD, Colombel J-F, Marteau P, Dore J: **Molecular inventory of faecal microflora in patients with Crohn's disease**. *FEMS Microbiol Ecol* 2004, **50**(1):25-36.

256. Beaven S, Abreu M: **Biomarkers in inflammatory bowel disease.** *Curr Opin Gastroenterol* 2004, **20**(4):318-327.

257. Neut C, Bulois P, Desreumaux P, Membree J-M, Lederman E, Gambiez L, Cortot A, Quandalle P, Kruiningen H, Colombel J-F: **Changes in the bacterial flora of the neoterminal ileum after ileocolonic resection for Crohn's disease**. *Am J Gastroenterol* 2002, **97**(4):939-946.

258. Ambrose N, Johnson M, Burdon D, Keighley M: **Incidence of pathogenic bacteria from mesenteric lymph nodes and ileal serosa during Crohn's disease surgery.** *Br J Surg* 1984 **71**(8):623-625.

259. Laffineur G, Lescut D, Vincent P, Quandalle P, Wurtz A, Colombel J: **Bacterial translocation in Crohn disease**. *Gastroenterol Clin Biol* 1992, **16**(10):777-781.

260. Ryan P, Kelly RG, Lee G, Collins JK, O'Sullivan GC, O'Connell J, Shanahan F: **Bacterial DNA within granulomas of patients with Crohn's disease -- Detection by laser capture microdissection and PCR**. *Am J Gastroenterol* 2004, **99**(8):1539-1543.

261. Schultsz C, Moussa M, van Ketel R, Tytgat GN, Dankert J: **Frequency of pathogenic and enteroadherent *Escherichia coli* in patients with inflammatory bowel disease and controls**. *J Clin Pathol* 1997, **50**(7):573-579.

262. Kotlowski R, Bernstein CN, Sepehri S, Krause DO: **High prevalence of *Escherichia coli* belonging to the B2+D phylogenetic group in inflammatory bowel disease**. *Gut* 2007, **56**(5):669-675.

263. Darfeuille-Michaud A, Neut C, Barnich N, Lederman E, Di Martino P, Desreumaux P, Gambiez L, Joly B, Cortot A, Colombel J-F: **Presence of adherent *Escherichia coli* strains in ileal mucosa of patients with Crohn's disease.** *Gastroenterology* 1998, **115**(6):1405-1413.

264. Nataro JP, Kaper JB: **Diarrheagenic *Escherichia coli***. *Clin Microbiol Rev* 1998, **11**(1):142-201.

265. Kaper JB, Nataro JP, Mobley HLT: **Pathogenic *Escherichia coli***. *Nat Rev Micro* 2004, **2**(2):123-140.

266. Chen L, Yang J, Yu J, Yao Z, Sun L, Shen Y, Jin Q: **VFDB: a reference database for bacterial virulence factors**. *Nucl Acids Res* 2005, **33**(suppl_1):D325-328.

267. Yang J, Chen L, Sun L, Yu J, Jin Q: **VFDB 2008 release: an enhanced web-based resource for comparative pathogenomics**. *Nucl Acids Res* 2008, **36**(suppl_1):D539-542.

268. Scaletsky ICA, Fabbricotti SH, Carvalho RLB, Nunes CR, Maranhao HS, Morais MB, Fagundes-Neto U: **Diffusely adherent *Escherichia coli* as a cause of acute diarrhea in young children in northeast Brazil: a case-control study**. *J Clin Microbiol* 2002, **40**(2):645-648.

269. Servin AL: **Pathogenesis of Afa/Dr Diffusely Adhering *Escherichia coli***. *Clin Microbiol Rev* 2005, **18**(2):264-292.

270. Foxman B: **Epidemiology of urinary tract infections: incidence, morbidity, and economic costs.** *Dis Mon* 2003 **49**(2):53-70.

271. Eto DS, Jones TA, Sundsbak JL, Mulvey MA: **Integrin-mediated host cell invasion by type 1-piliated uropathogenic *Escherichia coli***. *PLoS Pathog* 2007, **3**(7):e100.

272. Wiles TJ, Kulesus RR, Mulvey MA: **Origins and virulence mechanisms of uropathogenic *Escherichia coli***. *Exp Mol Pathol* 2008, **85**(1):11-19.

273. Martinez JJ, Mulvey MA, Schilling JD, Pinkner JS, Hultgren SJ: **Type 1 pilus-mediated bacterial invasion of bladder epithelial cells.** *EMBO J* 2000, **19**:2803-2812.

274. Wright KJ, Seed PC, Hultgren SJ: **Development of intracellular bacterial communities of uropathogenic *Escherichia coli* depends on type 1 pili**. *Cell Microbiol* 2007, **9**(9):2230-2241.

275. Boudeau J, Glasser A-L, Masseret E, Joly B, Darfeuille-Michaud A: **Invasive ability of an *Escherichia coli* strain isolated from the ileal mucosa of a patient with Crohn's disease**. *Infect Immun* 1999, **67**(9):4499-4509.

276. Boudeau J, Barnich N, Darfeuille-Michaud A: **Type 1 pili-mediated adherence of *Escherichia coli* strain LF82 isolated from Crohn's disease is involved in bacterial invasion of intestinal epithelial cells**. *Mol Microbiol* 2001, **39**(5):1272-1284.

277. Eaves-Pyles T, Allen CA, Taormina J, Swidsinski A, Tutt CB, Eric Jezek G, Islas-Islas M, Torres AG: ***Escherichia coli* isolated from a Crohn's disease patient adheres, invades, and induces inflammatory responses in polarized intestinal epithelial cells**. *Int J Med Microbiol* 2008, **298**(5-6):397-409.

278. Sasaki M, Sitaraman SV, Babbin BA, Gerner-Smidt P, Ribot EM, Garrett N, Alpern JA, Akyildiz A, Theiss AL, Nusrat A, Klapproth J-MA: **Invasive *Escherichia coli* are a feature of Crohn's disease**. *Lab Invest* 2007, **87**(10):1042-1054.

279. Glasser A-L, Boudeau J, Barnich N, Perruchot M-H, Colombel J-F, Darfeuille-Michaud A: **Adherent invasive *Escherichia coli* strains from patients with Crohn's disease survive and replicate within macrophages without inducing host cell death**. *Infect Immun* 2001, **69**(9):5529-5537.

280. Meconi S, Vercellone A, Levillain F, Payré B, Saati TA, Capilla F, Desreumaux P, Darfeuille-Michaud A, Altare F: **Adherent-invasive *Escherichia coli* isolated from Crohn's disease patients induce granulomas *in vitro***. *Cell Microbiol* 2007, **9**(5):1252-1261.

281. Barnich N, Boudeau J, Claret L, Darfeuille-Michaud A: **Regulatory and functional co-operation of flagella and type 1 pili in adhesive and invasive abilities of AIEC strain LF82 isolated from a patient with Crohn's disease**. *Mol Microbiol* 2003, **48**(3):781-794.

282. Barnich N, Bringer M-A, Claret L, Darfeuille-Michaud A: **Involvement of lipoprotein NlpI in the virulence of adherent invasive *Escherichia coli* strain LF82 isolated from a patient with Crohn's disease**. *Infect Immun* 2004, **72**(5):2484-2493.

283. Rolhion N, Barnich N, Claret L, Darfeuille-Michaud A: **Strong decrease in invasive ability and outer membrane vesicle release in Crohn's disease-associated adherent-invasive *Escherichia coli* strain LF82 with the *yfgL* gene deleted**. *J Bacteriol* 2005, **187**(7):2286-2296.

284. Rolhion N, Carvalho FA, Darfeuille-Michaud A: **OmpC and the sigma(E) regulatory pathway are involved in adhesion and invasion of the Crohn's**

disease-associated *Escherichia coli* strain LF82. *Mol Microbiol* 2007, **63**(6):1684-1700.

285. Bringer M-A, Rolhion N, Glasser A-L, Darfeuille-Michaud A: **The oxidoreductase DsbA plays a key role in the ability of the Crohn's disease-associated adherent-invasive *Escherichia coli* strain LF82 to resist macrophage killing**. *J Bacteriol* 2007, **189**(13):4860-4871.

286. Bringer M-A, Barnich N, Glasser A-L, Bardot O, Darfeuille-Michaud A: **HtrA stress protein is involved in intramacrophagic replication of adherent and invasive *Escherichia coli* strain LF82 isolated from a patient with Crohn's disease**. *Infect Immun* 2005, **73**(2):712-721.

287. Barnich N, Carvalho F, Glasser A, Darcha C, Jantscheff P, Allez M, Peeters H, Bommelaer G, Desreumaux P, Colombel J, Darfeuille-Michaud A: **CEACAM6 acts as a receptor for adherent-invasive *E. coli*, supporting ileal mucosa colonization in Crohn disease.** *J Clin Invest* 2007, **117**(6):1566-1574.

288. Darfeuille-Michaud A, Boudeau J, Bulois P, Neut C, Glasser A-L, Barnich N, Bringer M-A, Swidsinski A, Beaugerie L, Colombel J-F: **High prevalence of adherent-invasive *Escherichia coli* associated with ileal mucosa in Crohn's disease**. *Gastroenterology* 2004, **127**(2):412-421.

289. Martin HM, Campbell BJ, Hart CA, Mpofu C, Nayar M, Singh R, Englyst H, Williams HF, Rhodes JM: **Enhanced *Escherichia coli* adherence and invasion in Crohn's disease and colon cancer**. *Gastroenterology* 2004, **127**(1):80-93.

290. Glasser A-L, Darfeuille-Michaud A: **Abnormalities in the handling of intracellular bacteria in Crohn's disease: a link between infectious etiology and host genetic susceptibility**. *Arch Immunol Ther Exp* 2008, **56**(4):237-244.

291. Semrin G, Fishman DS, Bousvaros A, Zholudev A, Saunders AC, Correia CE, Nemeth E, Grand RJ, Weinstein DA: **Impaired intestinal iron absorption in Crohn's disease correlates with disease activity and markers of inflammation**. *Inflamm Bowel Dis* 2006, **12**(12):1101-1106.

292. Ren Z-P, Sällström J, Sundström C, Nistér M, Olsson Y: **Recovering DNA and optimizing PCR conditions from microdissected formalin-fixed and paraffin-embedded materials.** . *Pathobiology* 2000, **68**:215-217

293. Suau A, Rochet V, Sghir A, Gramet G, Brewaeys S, Sutren M, Rigottier-Gois L, Doré J: ***Fusobacterium prausnitzii* and related species represent a dominant group within the human fecal flora**. *Syst Appl Microbiol* 2001, **24**(1):139-145.

294. Wang R-F, Cao W-W, Cerniglia C-E: **Phylogenetic analysis of *Fusobacterium prausnitzii* based upon the 16S rRNA gene sequence and PCR confirmation**. *Int J Syst Bacteriol* 1996, **46**(1):341–343.

295. Sokol H, Seksik P, Furet JP, Firmesse O, Nion-Larmurier I, Beaugerie L, Cosnes J, Corthier G, Marteau P, Dore J: **Low counts of *Faecalibacterium prausnitzii* in colitis microbiota**. *Inflamm Bowel Dis* 2009, **DOI 10.1002/ibd.20903**.

296. Sokol H, Pigneur B, Watterlot L, Lakhdari O, Bermúdez-Humarán LG, Gratadoux J-J, Blugeon S, Bridonneau C, Furet J-P, Corthier G, Grangette C, Vasquez N, Pochart P, Trugnan G, Thomas G, Blottière HM, Doré J, Marteau P, Seksik P, Langella P: ***Faecalibacterium prausnitzii* is an anti-inflammatory commensal bacterium identified by gut microbiota analysis of Crohn disease patients**. *Proc Natl Acad Sci USA* 2008, **105**(43):16731-16736.

297. Duncan SH, Hold GL, Harmsen HJM, Stewart CS, Flint HJ: **Growth requirements and fermentation products of *Fusobacterium prausnitzii*, and a**

proposal to reclassify it as *Faecalibacterium prausnitzii* gen. nov., comb. nov. *Int J Syst Evol Microbiol* 2002, **52**(6):2141-2146.

298. Barcenilla A, Pryde SE, Martin JC, Duncan SH, Stewart CS, Henderson C, Flint HJ: **Phylogenetic relationships of butyrate-producing bacteria from the human gut**. *Appl Environ Microbiol* 2000, **66**(4):1654-1661.

299. Li M, Wang B, Zhang M, Rantalainen M, Wang S, Zhou H, Zhang Y, Shen J, Pang X, Zhang M, Wei H, Chen Y, Lu H, Zuo J, Su M, Qiu Y, Jia W, Xiao C, Smith LM, Yang S, Holmes E, Tang H, Zhao G, Nicholson JK, Li L, Zhao L: **Symbiotic gut microbes modulate human metabolic phenotypes**. *Proc Natl Acad Sci USA* 2008, **105**(6):2117-2122.

300. Pryde SE, Duncan SH, Hold GL, Stewart CS, Flint HJ: **The microbiology of butyrate formation in the human colon**. *FEMS Microbiol Lett* 2002, **217**(2):133-139.

301. Ramirez-Farias C, Slezak K, Fuller Z, Duncan A, Holtrop G, Louis P: **Effect of inulin on the human gut microbiota: stimulation of *Bifidobacterium adolescentis* and *Faecalibacterium prausnitzii***. *Br J Nutr* 2009, **101**(04):541-550.

302. Kleessen B, Schwarz S, Boehm A, Fuhrmann H, Richter A, Henle T, Krueger M: **Jerusalem artichoke and chicory inulin in bakery products affect faecal microbiota of healthy volunteers**. *Br J Nutr* 2007, **98**(03):540-549.

303. Scanlan PD, Shanahan F, Clune Y, Collins JK, O'Sullivan GC, O'Riordan M, Holmes E, Wang Y, Marchesi JR: **Culture-independent analysis of the gut microbiota in colorectal cancer and polyposis**. *Environ Microbiol* 2008, **10**(3):789-798.

304. Hayashi H, Sakamoto M, Kitahara M, Benno Y: **Diversity of the *Clostridium coccoides* group in human fecal microbiota as determined by 16S rRNA gene library**. *FEMS Microbiol Lett* 2006, **257**(2):202-207.

305. Collins MD, Lawson PA, Willems A, Cordoba JJ, Fernandez-Garayzabal J, Garcia P, Cai J, Hippe H, Farrow JAE: **The phylogeny of the genus *Clostridium*: Proposal of five new genera and eleven new species combinations**. *Int J Syst Bacteriol* 1994, **44**(4):812-826.

306. Finegold SM, Song Y, Liu C, Hecht DW, Summanen P, Könönen E, Allen SD: ***Clostridium clostridioforme*: a mixture of three clinically important species**. *Eur J Clin Microbiol Infect Dis* 2005, **24**(5):319-324.

307. Clayton EM, Rea MC, Shanahan F, Quigley EMM, Kiely B, Hill C, Ross RP: **The vexed relationship between *Clostridium difficile* and inflammatory bowel disease: An assessment of carriage in an outpatient setting among patients in remission**. *Am J Gastroenterol* 2009, **104**(5):1162-1169.

308. Namavar F, Theunissen EBM, Verweij-Van Vught AMJJ, Peerbooms PGH, Bal M, Hoitsma HFW, Maclaren DM: **Epidemiology of the *Bacteroides fragilis* group in the colonic flora in 10 patients with colonic cancer**. *J Med Microbiol* 1989, **29**(3):171-176.

309. Kuwahara T, Yamashita A, Hirakawa H, Nakayama H, Toh H, Okada N, Kuhara S, Hattori M, Hayashi T, Ohnishi Y: **Genomic analysis of *Bacteroides fragilis* reveals extensive DNA inversions regulating cell surface adaptation**. *Proc Natl Acad Sci USA* 2004, **101**(41):14919-14924.

310. Hoskins LC, Agustines M, McKee WB, Boulding ET, Krians M, Niedermeyer G: **Mucin degradation in human colon ecosystems**. *J Clin Invest* 1985, **75**:944-953.

311. Favier C, Neut C, Mizon C, Cortot A, Colombel JF, Mizon J: **Differentiation and identification of human faecal anaerobic bacteria producing [beta]-galactosidase (a new methodology)**. *J Microbiol Methods* 1996, **27**(1):25-31.

312. Macfarlane GT, Gibson GR: **Formation of glycoprotein degrading enzymes by *Bacteroides fragilis***. *FEMS Microbiol Lett* 1991, **77**(2-3):289-293.

313. Deplancke B, Vidal O, Ganessunker D, Donovan SM, Mackie RI, Gaskins HR: **Selective growth of mucolytic bacteria including *Clostridium perfringens* in a neonatal piglet model of total parenteral nutrition**. *Am J Clin Nutr* 2002, **76**:1117-1125.

314. Dwarakanath AD, Campbell BJ, Tsai HH, Sunderland D, Hart CA, Rhodes JM: **Faecal mucinase activity assessed in inflammatory bowel disease using ^{14}C threonine labelled mucin substrate.** *Gut* 1995 **37**(1):58-62.

315. Tsai HH, Dwarakanath AD, Hart CA, Milton JD, Rhodes JM: **Increased faecal mucin sulphatase activity in ulcerative colitis: a potential target for treatment.** *Gut* 1995 **36**(4):570-576.

316. El Yamani J, Mizon C, Capon C, Colombel JF, Fournet B, Cortot A, Mizon J: **Decreased faecal exoglycosidase activities identify a subset of patients with active Crohn's disease**. *Clin Sci* 1992, **83**(4):409-415.

317. Rhodes JM: **Unifying hypothesis for inflammatory bowel disease and associated colon cancer: sticking the pieces together with sugar.** *Lancet* 1996 **347**(8993):40-44.

318. Martinez-Medina M, Aldeguer X, Lopez-Siles M, González-Huix F, López-Oliu C, Dahbi G, Blanco JE, Blanco J, Garcia-Gil LJ, Darfeuille-Michaud A: **Molecular diversity of *Escherichia coli* in the human gut: new ecological evidence supporting the role of adherent-invasive *E. coli* (AIEC) in Crohn's disease.** *Inflamm Bowel Dis* 2009, **15**(6):872-882.

319. Caprilli R: **Why does Crohn's disease usually occur in terminal ileum?** *Med Gen Med* 2008, **2**(4):352-356.

320. Balamurugan R, Rajendiran E, George S, Samuel GV, Ramakrishna BS: **Real-time polymerase chain reaction quantification of specific butyrate-producing bacteria, *Desulfovibrio* and *Enterococcus faecalis* in the feces of patients with colorectal cancer**. *J Gastroenterol Hepatol* 2008, **23**(8pt1):1298-1303.

321. Moore WE, Moore LH: **Intestinal floras of populations that have a high risk of colon cancer**. *Appl Environ Microbiol* 1995, **61**(9):3202-3207.

322. Turnbaugh PJ, Hamady M, Yatsunenko T, Cantarel BL, Duncan A, Ley RE, Sogin ML, Jones WJ, Roe BA, Affourtit JP, Egholm M, Henrissat B, Heath AC, Knight R, Gordon JI: **A core gut microbiome in obese and lean twins**. *Nature* 2009, **457**(7228):480-484.

323. Malinen E, Rinttila T, Kajander K, Matto J, Kassinen A, Krogius L, Saarela M, Korpela R, Palva A: **Analysis of the fecal microbiota of irritable bowel syndrome patients and healthy controls with real-time PCR**. *Am J Gastroenterol* 2005, **100**(2):373-382.

324. Kassinen A, Krogius-Kurikka L, Mäkivuokko H, Rinttilä T, Paulin L, Corander J, Malinen E, Apajalahti J, Palva A: **The fecal microbiota of irritable bowel syndrome patients differs significantly from that of healthy subjects**. *Gastroenterology* 2007, **133**(1):24-33.

325. Collado MC, Calabuig M, Sanz Y: **Differences between the fecal microbiota of coeliac infants and healthy controls**. *Curr Issues Intest Microbiol* 2007, **8**(1):9-14.

326. Nadal I, Donant E, Ribes-Koninckx C, Calabuig M, Sanz Y: **Imbalance in the composition of the duodenal microbiota of children with coeliac disease**. *J Med Microbiol* 2007, **56**(12):1669-1674.

327. Sanchez E, Nadal I, Donat E, Ribes-Koninckx C, Calabuig M, Sanz Y: **Reduced diversity and increased virulence-gene carriage in intestinal enterobacteria of coeliac children**. *BMC Gastroenterol* 2008, **8**(1):50.

328. Sanz Y, Sánchez E, Marzotto M, Calabuig M, Torriani S, Dellaglio F: **Differences in faecal bacterial communities in coeliac and healthy children as detected by PCR and denaturing gradient gel electrophoresis**. *FEMS Immunol Med Microbiol* 2007, **51**(3):562-568.

329. Vaarala O, Atkinson MA, Neu J: **The "perfect storm" for type 1 diabetes: the complex interplay between intestinal microbiota, gut permeability, and mucosal immunity.** *Diabetes* 2008, **57**(10):2555-2562.

330. Lupp C, Robertson ML, Wickham ME, Sekirov I, Champion OL, Gaynor EC, Finlay BB: **Host-mediated inflammation disrupts the intestinal microbiota and promotes the overgrowth of *Enterobacteriaceae***. *Cell Host Microbe* 2007, **2**(2):119-129.

331. Dogan B, Klaessig S, Rishniw M, Almeida RA, Oliver SP, Simpson K, Schukken YH: **Adherent and invasive *Escherichia coli* are associated with persistent bovine mastitis**. *Vet Microbiol* 2006, **116**(4):270-282.

332. Simpson KW, Dogan B, Rishniw M, Goldstein RE, Klaessig S, McDonough PL, German AJ, Yates RM, Russell DG, Johnson SE, Berg DE, Harel J, Bruant G, McDonough SP, Schukken YH: **Adherent and invasive *Escherichia coli* is associated with granulomatous colitis in Boxer dogs**. *Infect Immun* 2006, **74**(8):4778-4792.

333. Bronowski C, Smith SL, Yokota K, Corkill JE, Martin HM, Campbell BJ, Rhodes JM, Hart CA, Winstanley C: **A subset of mucosa-associated *Escherichia coli* isolates from patients with colon cancer, but not Crohn's disease, share pathogenicity islands with urinary pathogenic *E. coli***. *Microbiology* 2008, **154**(2):571-583.

334. Barnich N, Darfeuille-Michaud A: **Adherent-invasive *Escherichia coli* and Crohn's disease.** *Curr Opin Gastroenterol* 2007, **23**(1):16-20.

335. Peeters H, Bogaert S, Laukens D, Rottiers P, Keyser FD, Darfeuille-Michaud A, Glasser AL, Elewaut D, Vos MD: ***CARD15* variants determine a disturbed early response of monocytes to adherent-invasive *Escherichia coli* strain LF82 in Crohn's disease**. *Int J Immunogenet* 2007, **34**(3):181-191.

336. Keita Å, Salim SY, Jiang T, Yang PC, Franzén L, Söderkvist P, Magnusson KE, Söderholm JD: **Increased uptake of non-pathogenic *E. coli* via the follicle-associated epithelium in longstanding ileal Crohn's disease**. *J Pathol* 2008, **215**(2):135-144.

337. Mpofu CM, Campbell BJ, Subramanian S, Marshall-Clarke S, Hart CA, Cross A, Roberts CL, McGoldrick A, Edwards SW, M. Rhodes J: **Microbial mannan inhibits bacterial killing by macrophages: A possible pathogenic mechanism for Crohn's disease**. *Gastroenterology* 2007, **133**(5):1487-1498.

338. Rodloff AC, Becker J, Blanchard DK, Klein TW, Hahn H, Friedman H: **Inhibition of macrophage phagocytosis by *Bacteroides fragilis* in vivo and in vitro**. *Infect Immun* 1986, **52**(2):488-492.

339. Nowrouzian F, Adlerberth I, Wold AE: **P fimbriae, capsule and aerobactin characterize colonic resident *Escherichia coli***. *Epidemiol Infect* 2001, **126**(1):11-18.

340. Nowrouzian FL, Adlerberth I, Wold AE: **Enhanced persistence in the colonic microbiota of *Escherichia coli* strains belonging to phylogenetic group B2: role of virulence factors and adherence to colonic cells**. *Microbes Infect* 2006, **8**(3):834-840.

341. Le Gall T, Clermont O, Gouriou S, Picard B, Nassif X, Denamur E, Tenaillon O: **Extraintestinal virulence is a coincidental by-product of commensalism in B2 phylogenetic group *Escherichia coli* strains**. *Mol Biol Evol* 2007, **24**(11):2373-2384.

342. Levin B: **The evolution and maintenance of virulence in microparasites.** *Emerging Infect Dis* 1996, **2**(2):93-102.

343. Costerton W, Veeh R, Shirtliff M, Pasmore M, Post C, Adami R, Ehrlich G: **The application of biofilm science to the study and control of chronic bacterial infections.** *J Clin Invest* 2003, **112**(10):1466-1477.

344. Hall-Stoodley L, Stoodley P: **Biofilm formation and dispersal and the transmission of human pathogens**. *Trends Microbiol* 2005, **13**(1):7-10.

345. Claret L, Miquel S, Vieille N, Ryjenkov DA, Gomelsky M, Darfeuille-Michaud A: **The flagellar sigma factor FliA regulates adhesion and invasion of Crohn's disease-associated *Escherichia coli* via a Cyclic dimeric GMP-dependent pathway**. *J Biol Chem* 2007, **282**(46):33275–33283.

346. Carvalho FA, Barnich N, Sauvanet P, Darcha C, Gelot A, Darfeuille-Michaud A: **Crohn's disease-associated *Escherichia coli* LF82 aggravates colitis in injured mouse colon via signaling by flagellin**. *Inflamm Bowel Dis* 2008, **14**(8):1051-1060.

347. Tieng V, Le Bouguénec C, du Merle L, Bertheau P, Desreumaux P, Janin A, Charron D, Toubert A: **Binding of *Escherichia coli* adhesin AfaE to CD55 triggers cell-surface expression of the MHC class I-related molecule MICA**. *Proc Natl Acad Sci USA* 2002, **99**(5):2977-2982.

348. Pizarro-Cerdá J, Cossart P: **Bacterial adhesion and entry into host cells**. *Cell* 2006, **124**(4):715-727.

Annexes

I. Annexed tables

Annexed table 1. Partial 16S rRNA gene sequences obtained in this work and best BLAST matches against GenBank.

Sequence ID	Nearest sequence	Similarity (%)	Source	Nearest known species (Acc. Num)	Similarity (%)
C0201	-			*Ruminococcus obeum* (AY169419)	99
C0301	AF371512	99	pig gastrointestinal tract	*Eubacterium biforme* (M59230)	98
C0302	AY984355	98	human intestinal tract	*Ruminococcus obeum* (AY169419)	98
C0303	AF153858	97	feces	*Ruminococcus lactaris* (L76602)	97
C0304	AY471654	97	stools from patients with Crohn's disease	*Ruminococcus obeum* (AY169419)	96
C0305	AF530354	98	stools from patients with Crohn's disease	*Faecalibacterium prausnitzii* (AY169429)	98
C0401	AY985869	99	human intestinal tract	*Faecalibacterium prausnitzii* (X85022)	98
C0402	AY986207	99	human intestinal tract	*Faecalibacterium prausnitzii* (X85022)	98
C0403	AY983968	98	human intestinal tract	*Catenibacterium mitsuokai* (AB030222)	97
C0404	AY919925	96	feces	*Rikenella microfusus* ATCC 29728 (L16498)	83
C0405	-			*Atopobium oviles* (AJ251324)	100
C0501	AF371512	97	pig gastrointestinal tract	*Eubacterium biforme* (M59230)	95
C0502	AY983727	99	human intestinal tract	*Clostridium cocleatum* (AF028350)	97
C0503	-			*Escherichia coli* RREC I (AF527827)	99
C0504	-			*Escherichia coli* RREC III (AF527825)	99
C0601	AF153858.1	99	feces	*Ruminococcus lactaris* (L76602)	99
C0701	AY985177	99	human intestinal tract	*Eubacterium biforme* (M59230)	98
C0702	AY985486	100	human intestinal tract	*Faecalibacterium prausnitzii* (X85022)	99
C0703	AY986207	93	human intestinal tract	*Faecalibacterium prausnitzii* (X85022)	92
C0704	-			*Bacteroides massiliensis* B84634 (AY126616)	96
C0801	AY986207	99	human intestinal tract	*Faecalibacterium prausnitzii* (X85022)	98
C0901	AY985948	99	human intestinal tract	*Ruminococcus torques* (L76604)	96
C0902	AY592220	100	root nodules of legume species	*Ruminococcus torques* (L76604)	96
C0903	AY592220	100	anaerobic methane oxidizers from Kazan mud volcano (eastern Mediterranean)	*Ruminococcus torques* (L76604)	96
C1001	AY684430	99	human intestinal tract	*Fusobacterium necrogenes* (X55408)	98
C1201	AY494687	100	Salmonid gill	*Bacteroides vulgatus* (AB050111)	99
C1401	AF153858	99	feces	*Ruminococcus lactaris* (L76602)	99
C1402	AY985869	100	human intestinal tract	*Faecalibacterium prausnitzii* (X85022)	99
C1403	AY986349	99	human intestinal tract	*Faecalibacterium prausnitzii* (X85022)	98
C1501	AY986341	100	human intestinal tract	*Bacteroides vulgatus* (AM042696)	99
CD0101	AY452007	99	intestinal mucosa	*Ruminococcus torques* (L76604)	98
CD0102	-			*Escherichia coli* RREC III (AF527825)	99
CD0103	-			*Enterococcus faecalis* (AY850358)	98
CD0201	AY305314	100	human colon	*Clostridium boltei* 16351 (AJ508452	97
CD0202	Y10028	100	rumen of red deer	*Clostridium clostridioforme* (M59089)	99
CD0203	-			*Ruminococcus gnavus* (X94967)	99
CD0204	-			*Ruminococcus gnavus* (X94967)	98
CD0205	AJ408999	98	human colon	*Ruminococcus schinkii* (X94964)	97
CD0206	-			*Ruminococcus schinkii* (X94964)	97
CD0207	-			*Corynebacterium simulans* UCL557 (AJ012838)	100
CD0208	-			*Corynebacterium simulans* UCL557 (AJ012838)	100
CD0209	-			*Corynebacterium simulans* UCL557 (AJ012838)	100
CD0210	-			*Corynebacterium simulans* UCL557 (AJ012838)	100
CD0211	AJ315487	96	human colon	*Clostridium hylemonae* (AB117570)	95
CD0302	AY471702.1	98	stools from patients with Crohn's disease	*Faecalibacterium prausnitzii* (AJ270469)	97
CD0303	-			*Escherichia coli* RREC III (AF527825)	99
CD0304	-			*Escherichia coli* RREC III (AF527825)	98
CD0305	AY916305	97	human intestinal tract	*Faecalibacterium prausnitzii* (X85022)	96
CD0401	AY452007	99	intestinal mucosa	*Ruminococcus torques* (L76604)	98
CD0402	AY452007	100	gut microbiota	*Ruminococcus torques* (L76604)	99
CD0403	AY916138	99	human intestinal tract	*Ruminococcus obeum* (X85101)	99
CD0404	AY916138	94	human intestinal tract	*Ruminococcus obeum* (X85101)	98
CD0405	AF153858	99	feces	*Ruminococcus lactaris* (L76602)	99
CD0406	AF153858	96	feces	*Ruminococcus lactaris* (L76602)	96
CD0407	AJ315487	99	human colon	*Clostridium hylemonae* (AB117570)	98
CD0408	AJ315487	99	human colon	*Clostridium hylemonae* CT-35 (AB117570)	98
CD0501	AY916305	97	human intestinal tract	*Faecalibacterium prausnitzii* (AJ270469)	96
CD0502	-			*Ruminococcus gnavus* (X94967)	97
CD0504	-			*Alistipes finegoldii* 4401054 (AY643082)	99
CD0505	AF132260	95	human gut	*Eubacterium ramulus* (AJ011522)	95
CD0506	AF132270.1	95	human gut	*Clostridium saccharolyticum* (Y18185)	95
CD0507	AY986226	97	human intestinal tract	*Bacteroides putredinis* ATCC 29800/*Alistipes putredinis* (L16497)	94
CD0508	AY452007	99	intestinal mucosa	*Ruminococcus torques* (L76604)	98
CD0509	-			*Ruminococcus gnavus* (X94967)	99
CD0601	AY920077	98	human feces	*Clostridium spiriforme* (X75908)	92
CD0602	AY920077	100	human feces	*Clostridium spiriforme* (X75908)	93
CD0603	AY986074	100	human intestinal tract	*Clostridium spiriforme* (X75908)	94
CD0604	AY920077	99	human feces	*Clostridium spiriforme* (X75908)	94
CD0605	AY986341	94	human intestinal tract	*Bacteroides vulgatus* (AB050111)	94
CD0606	AY452007	99	intestinal mucosa	*Ruminococcus torques* (L76604)	98
CD0607	AY452003	100	intestinal mucosa	*Ruminococcus obeum* (X85101)	96
CD0608	-			*Ruminococcus gnavus* (X94967)	100
CD0702	AY985751	99	human intestinal tract	*Bacteroides caccae* ATCC 43185T (X83951)	99
CD0703	AY985751	98	human intestinal tract	*B. caccae* ATCC 43185T (X83951)	98

Annexed table 1. (*Continued*)

CD0704	AY981791	96	human intestinal tract	*Faecalibacterium prausnitzii* (AJ270469)	95
CD0705	AY985177	99	human intestinal tract	*Eubacterium biforme* . (M59230)	98
CD0706	AY985177	99	human intestinal tract	*Eubacterium biforme* (M59230)	98
CD0707	AY985177	100	human intestinal tract	*Eubacterium biforme* (M59230)	98
CD0708	AY916138	99	human intestinal tract	*Ruminococcus obeum* (X85101)	99
CD0709	AY984355	98	human intestinal tract	*Ruminococcus obeum* (AY169419)	98
CD0710	AY916138	98	human intestinal tract	*Ruminococcus obeum* (X85101)	97
CD0711	AF153858	98	feces	*Ruminococcus lactaris* (L76602)	98
CD0712	AF153858	98	feces	*Ruminococcus lactaris* (L76602)	97
CD0713	AY983968	99	human intestinal tract	*Catenibacterium mitsuokai* (AB030222)	98
CD0714	AY983968	98	human intestinal tract	*Catenibacterium mitsuokai* (AB030222)	97
CD0715	AY983968	99	human intestinal tract	*Catenibacterium mitsuokai*:JCM 10607 (AB030222)	98
CD0716	AY916138	97	human intestinal tract	*Ruminococcus obeum* (X85101)	96
CD0717	AY388411	100	human feces	*Collinsella aerofaciens* H818 (AJ245920)	99
CD0801	AY977866.1	95	human intestinal tract	*Ruminococcus obeum* (AY169419)	95
CD0802	AY684430.2	100	human intestinal tract	*Fusobacterium necrogenes* (X55408)	98
CD0803	AY684430	100	human intestinal tract	*Fusobacterium mortiferum* (X55414)	100
CD0804	AY452007.1	100	intestinal mucosa	*Ruminococcus torques* (L76604)	99
CD0805	-			*Ruminococcus gnavus* (X94967)	100
CD1002	AY452007	98	intestinal mucosa	*Ruminococcus torques* (L76604)	97
CD1003	-			*Clostridium perfringens* (Y12669)	99
CD1004	AY916216	97	human intestinal tract	*Ruminococcus obeum* (X85101)	97
CD1005	-			*Klebsiella pneumoniae* (AY292865)	100
CD1007	-			*Staphylococcus epidermidis* KL-096 (AY030342.1)	100
CD1101	AY985581	100	human intestinal tract	*Bacteroides vulgatus* (AB050111)	99
CD1102	-			*Fusobacterium varium* (M58686)	99
CD1103	AF530331	99	stools from patients with Crohn's disease	*Clostridium nexile* (X73443)	99
CD1104	AJ315487	99	human colon	*Clostridium hylemonae*:CT-35 (AB117570)	99
CD1105	AJ315487	97	human colonic samples	*Ruminococcus gnavus* (X94967)	95
CD1201	-			*Proteus mirabilis* (AJ301682)	99
CD1202	-			*Escherichia coli* RREC III (AF527825)	99
CD1203	-			*Enterococcus faecalis* SFL (AY850358)	100
CD1301	AY986341.1	99	human intestinal tract	*Bacteroides vulgatus* (AB050111)	99
CD1401	AY986349	99	human intestinal tract	*Faecalibacterium prausnitzii* (AJ270469)	99
CD1402	-			*Haemophilus parainfluenzae* (AY362908)	99
CD1403	-			*Ruminococcus gnavus* (X94967)	100
CD1501	-			*Escherichia coli* RREC I (AF527827)	99
CD1502	-			*Escherichia coli* RREC III (AF527825)	100
CD1503	AY531211	98	root nodules of legume species	*Enterobacter cloacae* (Y17665)	98
CD1601	AY305313	99	human colon	*Eubacterium cylindroides* ATCC27803 (L34617)	99
CD1602	AY305313	99	butyrate-producing bacterium SM7/11	*Eubacterium cylindroides* (L34617)	99
CD1603	DQ057466	97	ileum and cecum of broiler chickens	*Faecalibacterium prausnitzii* (AY169429)	92
CD1604	-	100		*Ruminococcus gnavus* (X94967)	100
CD1701	AF530337	100	stools from patients with Crohn's Disease	*Clostridium cocleatum* DSM 1551 (Y18188.1)	100
CD1801	AY452007	94	intestinal mucosa	*Ruminococcus lactaris* (L76602)	93
CD1802	AY916258	95	human intestinal tract	*Clostridium hathewayi* (AY552788)	93
CD1901	AY985869	95	human intestinal tract	*Faecalibacterium prausnitzii* (X85022)	95
CD1902	AY986207	99	human intestinal tract	*Faecalibacterium prausnitzii* (AJ270470)	99
CD1903	AY388410	94	human cecum and feces	*Eubacterium rectale* S2Ss2/2 (AY804152)	94
CD1904	AY985152	100	human intestinal tract	*Eubacterium rectale* S2Ss2/2 (AY804152)	100
CD1905	AF371609.1	100	pig gastrointestinal tract	*Eubacterium rectale* S2Ss2/2 (AY804152.1)	94
CD1906	AY985152	100	human intestinal tract	*Eubacterium rectale* S2Ss2/2 (AY804152.1)	100
CD1907	AF371609.1	100	pig gastrointestinal tract	*Eubacterium rectale* S2Ss2/2 (AY804152.1)	100
CD1908	AF371609	95	pig gastrointestinal tract	*Roseburia faecalis* M88/1 (AY804150)	95
CD1909	AY986276	98	human intestinal tract	*Faecalibacterium prausnitzii* (AJ270469)	98
CD1910	AY985869	98	human intestinal tract	*Faecalibacterium prausnitzii* (X85022)	98
IC0101	AY956677	100	cotton rhizosphere anb bulk soil from monoculture	*Escherichia coli* RREC III (AF527825)	100
UC0101	AJ315487	94	human colon	*Clostridium hylemonae* (AB117570)	94
UC0102	AY452007	96	intestinal mucosa	*Ruminococcus torques* (L76604)	95
UC0102	AY986207	99	human intestinal tract	*Faecalibacterium prausnitzii* (AJ270470)	98
UC0103	AY452007	96	intestinal mucosa	*Ruminococcus torques* (L76604)	95
UC0201	-			*Bacteroides thetaiotaomicron* 8736 (AY895203)	98
UC0203	-			*Bacteroides thetaiotaomicron* 8736 (AY895203)	95
UC0204	-			*Clostridium boltei* 16351 (AJ508452.1)	100
UC0205	AY983739.1	97	human intestinal tract	*Dorea longicatena* (AJ132842)	100
UC0206	AY985642.1	100	human intestinal tract	*Dorea longicatena* (AJ132842)	100
UC0208	AY452007.1	100	intestinal mucosa	*Ruminococcus schinkii* Bie 41 (X94964)	100

BLAST performed in June 2006.
Nomenclature bands: CD/UC/IC/C Diagnose patient followed of two digits for Id patient and two more d for Id band. Example: CD0203, third band of CD patient number 2.

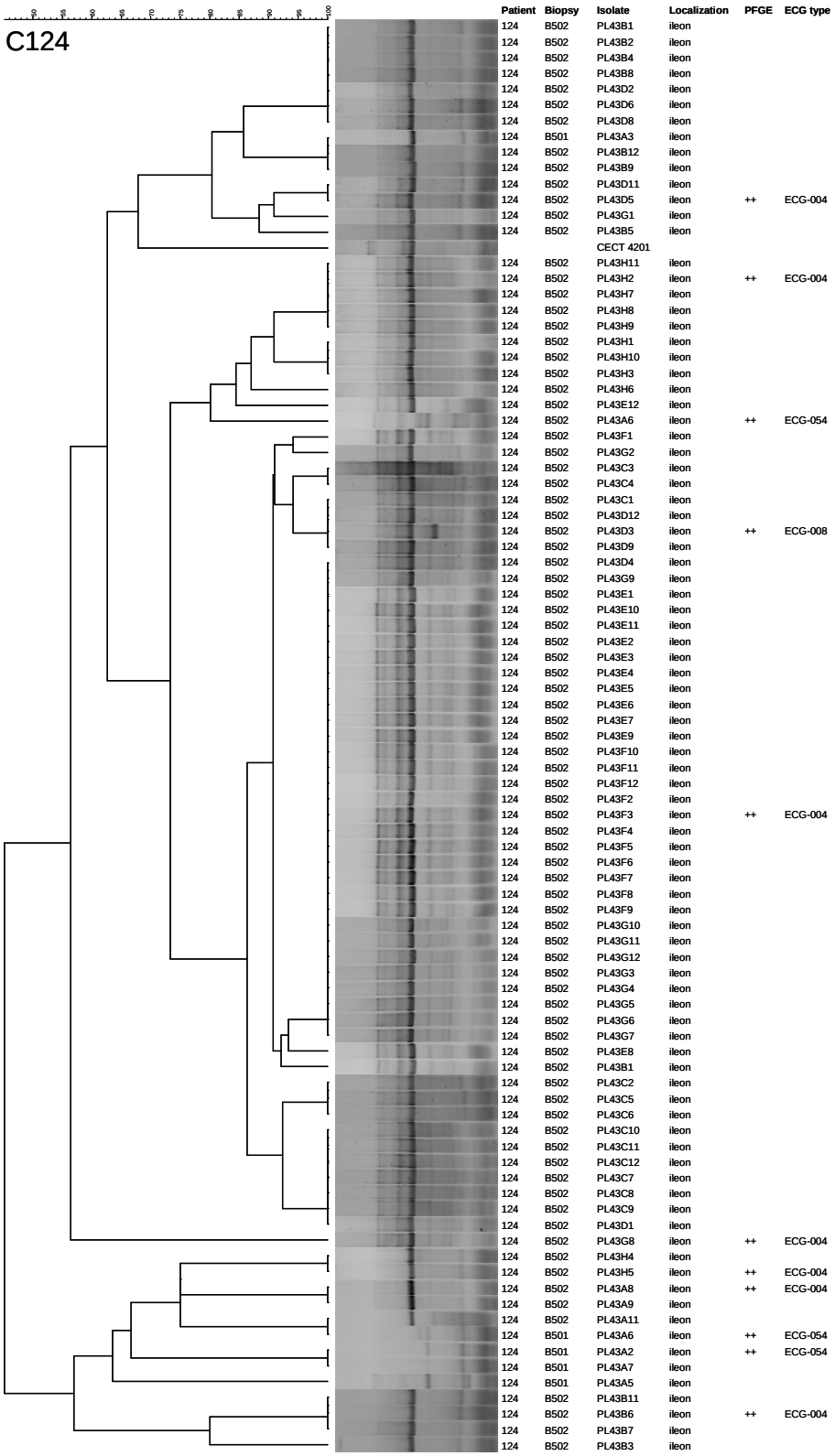
Annexed table 2. Mucosa-associated *E. coli* subtype richness, relative abundance and diversity

	E. coli subtype abundance												Total *E. coli* isolates	RICHNESS			*E. coli* subtype relative abundance										DIVERSITY		
ID patient	N	#	N	#	N	#	N	#	N	#	N	#		R	Mean	SD	N	#	N	#	N	#	N	#	N	#	H'	Mean	SD
C69	95	41											95	1	2.000	1.044	1.000	41									0.000	3.116	3.127
C70	1	22	21	25	21	60							43	3			0.023	22	0.488	25	0.488	60					5.195		
C71	33	24	9	25	42	60	1	12					85	4			0.388	24	0.106	25	0.494	60	0.012	12			8.339		
C80	116	43											116	1			1.000	43									0.000		
C81	11	46											11	1			1.000	46									0.000		
C82	10	16	1	37									11	2			0.909	16	0.091	37							2.493		
C83	26	17											26	1			1.000	17									0.000		
C84	94	49											94	1			1.000	49									0.000		
C119	74	23	15	31	1	32							90	3			0.822	23	0.167	31	0.011	32					6.487		
C127	2	13	88	55									90	2			0.022	13	0.978	55							3.829		
C124	77	4	1	8	9	54							87	3			0.885	4	0.011	8	0.103	54					6.857		
C125	64	4	1	36									65	2			0.985	4	0.015	36							4.190		
CD53	160	1	1	2	15	9							176	3	2.300	1.494	0.909	1	0.006	2	0.085	9					7.728	4.513	5.347
CD55	34	18											34	1			1.000	34									0.000		
CD61	1	9	112	65	1	47							114	3			0.009	9	0.982	65	0.009	47					9.490		
CD79	2	19	21	42									23	2			0.087	19	0.913	42							2.533		
CD109	47	21	6	34									53	2			0.887	21	0.113	34							2.299		
CD113	84	15											84	1			1.000	15									0.000		
CD110	9	57	1	58	7	53	80	56	5	63	4	59	106	6			0.085	57	0.009	58	0.066	53	0.755	56	0.047	63	0.038	59	16.460		
CD111	160	28											160	1			1.000	28									0.000		
CD112	167	26	1	27									168	2			0.994	26	0.006	27							5.130		
CD72	25	5	13	64									38	2			0.658	5	0.342	64							1.491		

#, Id subtype (ECGx); N, number of isolates of a certain subtype
H', Shannon Index, calculated as = - (LN (relative abundance subtype 1)+ LN (relative abundance subtype 2)+...)
R, richness: number of different *E. coli* subtypes per patient

II. Annexed figures

Annexed figure 1. First screening of the clonality of *E. coli* isolates obtained from a single subject.



This figure is an example of the dendrogram resulting from the IS3-based Rep-PCR profiles of these *E. coli* isolates. Those for which the clonality was further confirmed by PFGE are indicated with ++.

III. Glossary (words marked with * in the text)

Abscess	A localised collection of pus caused by suppuration buried in tissue, organs or confined spaces. Usually due to an infective process.
Ankylosing spondylitis	A polyarthritis involving the spine, which is characterised by progressive, painful stiffening of the joints and ligaments. It almost exclusively affects young men.
Arthralgia	Severe pain, mainly on flexion of the joints of the lower extremities.
Arthritis	An inflammatory condition that affects joints. Can be infective, autoimmune, or traumatic in origin.
Autophagy	A cell-autonomous process involved in the degradation of intracellular pathogens, antigen processing, regulation of cell signalling and regulation of T-cell homeostasis.
Concordance	The occurrence of the same trait in both members of a pair of twins. Concordance might occur for diseases as well as for behaviours such as smoking.
Crypts	Tubular invaginations of the intestinal epithelium. At the base of the crypts, there are Paneth cells, which produce bactericidal defensins, and stem cells, which continuously divide and are the source of all intestinal epithelial cells.
Defensins	A family of proteins exhibiting bactericidal properties. They are secreted by immune cells (particularly neutrophils), intestinal Paneth cells and epithelial cells.
Edema	The presence of abnormally large amounts of fluid in the intercellular tissue spaces of the body.
Episcleritis	Inflammation of the thin membrane which coats the sclera (white of the eye).
Erythema nodosum	A disorder characterised by the formation of tender, red nodules on the fronts of the legs. Erythema nodosum primarily affects women and has been associated with certain infections, particular drugs, and certain diseases such as leukaemia, sarcoidosis, rheumatic fever, ulcerative colitis and with pregnancy.
Fibrosis	The formation of excessive fibrous tissue.
Fistula	An abnormal passage or communication, usually between two internal organs or leading from an internal organ to the surface of the body.
Granuloma	A collection of epithelial cells, macrophages and lymphocytes, usually T cells.
Incidence	In epidemiology, incidence rate is the number of new cases that acquire a disease per unit of person and time.
Ischemia	A low oxygen state usually due to obstruction of the arterial blood supply or inadequate blood flow leading to hypoxia in the tissue.
Meta-analysis	"A statistical approach that combines results from multiple related studies to define a composite effect. When applied to genome-wide association studies, increased power to identify more modest association effects accrues" [47].
Microbiome	Genomes of human microbial symbionts.

Glossary (*Continued*)

Prevalence	In epidemiology, total number of cases with a certain disease in the population at a given time.
Pulsotype	Term used to define specific profiles obtained by pulsed field gel electrophoresis, characteristic of each clone.
Pyoderma gangrenosum	Uncommon ulcerative cutaneous condition of uncertain etiology. The lesion(s) usually begin as a soft nodule on the skin which proceeds to ulcerate. The ulcer enlarges and the skin at the edge is purple-red. This condition is associated with several other diseases, some of which are ulcerative colitis, Crohn's disease, rheumatoid arthritis, leukaemia, and cryoglobulinaemia.
Sacroiliitis	Inflammation of the sacroiliac joint. Can be caused by multiple factors and can present in a variety of different forms.
Scleritis	Refers to any inflammation of the sclera including episcleritis*, a benign condition affecting only the episclera, which is generally short-lived and easily treated. Classic scleritis, on the other hand, affects deeper tissue and is characterised by higher rates of visual acuity loss and even mortality, particularly in its necrotising form. Its characteristic symptom is severe and general head pain. Etiology is unknown but is thought to involve a local immune response.
Seropathotype	Term used to define the combination of several typing methods, usually including serotyping and virulence gene typing.
Siderophore	Low molecular weight molecule secreted by certain advantaged bacterial species that have a high affinity for ferric (Fe^{+3}) iron, which is insoluble as a free cation. Bacteria retrieve iron-bound siderophores through receptors and incorporate them into the cytoplasm.
Stenosis	Narrowing or stricture of a duct or canal.
Tenesmus	Straining, especially ineffectual and painful, during defecation.
Toxic megacolon	Complication of IBDs, mainly UC, characterised by a very dilated colon, accompanied by abdominal distension, and sometimes fever, abdominal pain, or shock. High risk of death.
Transmural	Extending through, or affecting the entire thickness of, the wall of an organ.
Uveitis	An inflammation of part or all of the uvea, which is the tunic that surrounds the eye and is made up of the iris, the ciliary body and the choroid.

Acknowledgements

This work was mainly funded by the **Spanish Ministry of Education and Science** (SAF2006-00414; 2006) and by the Foundation **La Marató de TV3** (2002). The **University of Girona** supported two research fellowships (2005 and 2006).

≡≡≡

I would like to thank the expert reviewers that have accepted to revise and evaluate this thesis as well as the members of the examining tribunal for the time you have spent reading this work.

Tota tesi és fruit de la participació de diverses persones que hi intervenen a diferents nivells. En aquesta tesi, en particular, han estat moltes les persones que m'han ajudat, per a mi totes de gran valor. És per això que vull agrair la vostra contribució en aquest treball:

Als **titulars del l'Àrea de Microbiologia**, per oferir-me una contractació digna com a professora associada a temps complert tant bon punt vaig acabar la llicenciatura.

A en **Jesús**, per confiar amb mi i donar-me l'oportunitat de treballar en aquest projecte que, a banda de ser un gran repte, m'ha motivat des del primer dia.

Als metges i infermeres del servei de digestiu de l'hospital Dr. Josep Trueta, l'hospital Santa Caterina i la Clínica de Girona, en especial als doctors **Xavier Aldeguer**, **Ferran Gonzálex Huix**, **Carles López Oliu** i **Doroteo Acero**. Per la col·laboració establerta aquests anys, primordial per a poder dur a terme aquest treball. Seria molt satisfactori veure que en un futur aquesta col·laboració es manté i que la recerca resultant sigui profitosa per a les persones que pateixen la malaltia de Crohn o d'altres malalties intestinals.

Al laboratori de la Dra. Arlette Darfeuille Michaud, on vaig aprendre a identificar el patovar AIEC, pas necessari per a la realització dels capítols 2 i 3 d'aquest treball. **Arlette**, je remercie très sincèrement ton apport à ce travail. Mon stage au Laboratoire de Pathogénie Bactérienne Intestinale à été très agréable du début à la fin. Tu m'as donné des bons conseils scientifiques et personales. **Nath**, merci pour un milliard the choses, tes conseils sur les manips, tes gâteaux, ta sympathie, et pour la meilleur crêpe de Londres !

Anne-Lise, pour ton très important aide avant arriver à Clermont. Tu resteras dans ma mémoire. **Nico**, tu es qui m'a montré comment travailler avec cellules eucaryotes, chose indispensable pour faire cette thèse. **Pedro**, sans commentaires… cua-cua ! **Laurent**, **Sylvie**, **Fred**, **Adeline**, **Vincent** et **Valérie** pour votre chaleureux accueil.

Al Laboratorio de Referencia de *Escherichia coli*, dirigido por el Dr. **Jorge Blanco**, quienes han realizado el serotipado y genotipado de las cepas de *E. coli*. Jorge, gracias por el apoyo y la confianza depositada. **Azucena**, gracias por tu implicación en el trabajo que en esta tesis se incluye en el capítulo 3.2, y por las maravillosas tapas de Lugo, ¡me quedo con el pulpo!

Al Departamento de Microbiología Médica y Quimioterapia Antimicrobiana de la Fundación Jiménez Díaz-Capio, quienes han realizado los ensayos de formación de biofilms. En especial a **Francisco Soriano**, quien además de tomar parte en la coordinación de esta parte del trabajo ha revisado detalladamente el manuscrito incluido en este trabajo como capítulo 3.1.

A la **Mireia**, sempre al peu del canó preparada per a fer el que calgui. No cal que et digui que estic molt contenta de tenir-te al costat i que et desitjo el millor perquè t'ho mereixes.

To the Laboratory of Veterinary Molecular Microbiology (University of Bristol) headed by Dr. **Vazquez-Boland**, where I learned the pulsed field gel electrophoresis assay under the supervision of **Debbi Lewis**. There I met wonderful people: **Alain**, **Cristina**, **Lizeth**, **Magda**…

Al grup de Bioquímica i Biologia Molecular de la Universitat de Girona, en especial a la Dra. **Maria Vilanova** i la **Jess**, per permetre'm utilitzar el seu espai de cultiu cel·lular i per acollir-m'hi tant bé.

A tota la gent del laboratori del Departament de Salut Pública de la Generalitat (delegació de Girona), on he realitzat bona part dels gels d'electroforesi de camp polsant, i en especial a la Dra. **Joana Pardos**, qui m'ha obert les portes sempre que ha calgut.

A la **Marie-Hélène Nicolas-Chanoine** (Service de Microbiologie, Hôpital AP-HP Beaujon, Clichy; INSERM U773, Faculté de Médecine D. Diderot, Université de Paris) i l'**Stéphane Bonacorsi** (Laboratoire d'études de Génétique Bactérienne dans les Infections de l'enfant, Faculté de Médecine D. Diderot, Université de Paris; Service de Microbiologie AP-HP Hôpital Robert Debré), per cedir-nos algunes soques ExPEC necessàries per al capítol 3.2 i a la vegada per haver revisat el treball inclòs en aquesta part de la tesi.

A en **Miguelito**, per ajudar-me amb l'estadística (cosa que ha estat molt important), també, per la seva alegria i per la seva música!

A la **Geno**, qui ha seguit de prop aquest treball, m'ha animat quan ha fet falta i s'ha preocupat del meu benestar en tot moment.

A en **Bo**, per escoltar-me i fer-me escoltar.

A en **Catxo**, per les teves preguntes i per les teves respostes.

A en **Xevi T**, amb qui he compartit molts dies de laboratori, despatx i esbarjo plens de bones estones.

A la **Laia Calvó**, que m'ha servit de referent en molts aspectes.

A en **Gela**, de qui encara he d'aprendre moltes coses, l'**Olaya**, intel·ligent i dolça, l'**Anna**, optimisme i implicació, l'**Àlex**, tendre i divertit, en **Marc**, extravagant i atent, la **Laia M**, gairebé la meva ànima bessona, per la vostra amistat.

A la **Núria** i la **Tere**, per la seva paciència, ja que han patit i pateixen la ocupació del −80ºC del laboratori de patògens amb els milers d'aïllats d'*E. coli*. Hem d'aconseguir un laboratori més digne!!

A l'**Olga**, per fer-me el *passe* de cèl·lules aquell dia que jo estava fora, pels teus consells sobre la tesi, i per la informació sobre l'SSH, entre d'altres.

A la **Noemí**, la **Goretti**, la **Maria**, la **Gis**, en **Lluís**, en **Vicenç**, en **Martí** i la resta de *Coffee Girls*, pels magnífics anys que hem compartit a la UdP.

A la **Noe**, per ser SEMPRE al meu costat, per la teva comprensió i estimació. Tots els que la coneixem sabem el que val!

A mis padres **Felipe** y **Margarita**, a mis tíos –en especial mi tío **Salva**, el diseñador de la portada de este libro–, a mis hermanos **David**, **Guillermo**, **Felipe**, **Bea** y **Cris**, y a mi abuela **Beatriz**, por su confianza y apoyo.

A en **Narcís**, qui ha viscut més de prop aquesta tesi, per la llibertat que m'has donat i pel que t'he fet patir.

Sempre vostra.

Grup de Microbiologia de la Malaltia Intestinal
Facultat de Ciències, Universitat de Girona

Campus de Montilivi s/n, E-17071, Girona
Tel.: +34 972 41 82 61 • Fax: +34 972 41 81 50
marga.martinez@udg.edu