

We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900

Open access books available

186,000

International authors and editors

200M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com



Evolutionary Insights into the “Population-Specificity” of the Genetic Factors Associated with Inflammatory Bowel Diseases

Shigeki Nakagome¹ and Hiroki Oota²

¹*The Institute of Statistical Mathematics,*

²*Kitasato University School of Medicine*

Japan

1. Introduction

Inflammatory bowel disease (IBD) – a generic term for Crohn’s disease (CD) and ulcerative colitis (UC) – represents great risks for inheritance, such as high sibling recurrent risks (λ s in CD and UC: 20-35 and 8-15) and high monozygotic concordance rates (50% to 60%). Recent genome-wide association studies (GWAS) have successfully identified many causative genes (risk variants) for IBD. These discoveries highlight the importance of autophagy and innate immunity as determinants of dysregulated host-commensal bacteria interactions in the intestine. Some of the risk-variants are shared between CD and UC, indicating both similarities and differences in their etiologies. Most of the IBD risk-variants, however, have been reported in European-ancestry populations, and these associations are not always reproduced in non-European ancestry populations, suggesting a population-specific susceptibility to IBD. In fact, some of the IBD risk variants showed no association with CD and UC in East Asians. Hence, IBD’s causative factors should be defined as genetically different between Europeans and East Asians; it is essential that we expand GWAS into non-European populations so as to unravel the causes of its population-specific susceptibility and identify IBD risk variants in each geographic population. It is widely an accepted idea that the risk variants of complex diseases, including IBD, are retained not only in patients, but also in healthy controls. Because the risk variants spread prevalently into geographically diverse populations, it is plausible that their origins could be dated before *Homo sapiens*’ expansion out of Africa; that is, the fate of risk variants must be affected by human population history, such as demography, migration and natural selection. Here, we suggest that an evolutionary perspective of IBD genomics could provide essential clues for resolving significant questions: (1) Do the risk-variants have similar allele frequencies in different populations? (2) How are the risk-variants prevalent in human populations? (3) What factors can cause the inconsistency of GWAS-results across geographic populations? In this chapter, we first provide an overview of the evolutionary characteristics of disease-causative variants in Mendelian diseases and complex diseases. Secondly, we review the recent progress of IBD genetic/genomic studies among both European and East Asian populations. Finally, we discuss the evolutionary consequences of population-specific susceptibility to IBD and the importance of its use for diverse human populations in the future of GWAS.

2. Genetic studies of IBD and population-specific susceptibility

IBDs are complex diseases in which multiple genetic and environmental factors are likely to contribute to pathogenesis. Human genomics studies have concentrated their attention on the identification of the genetic variants underlying these diseases over the past decade, since none of the complex diseases are inherited in a simple Mendelian fashion. The genetic characteristics of variants are significantly different between Mendelian (monogenic) diseases and complex (multi-factorial) diseases, both of which constitute the basis of genetic mapping strategies (e.g., linkage analysis and case-control association studies) for the revelation of their role in the aetiology of clinically defined phenotypes. Disease alleles are a specific subset of all the genetic variants present in human populations, and the origin of diseases can be addressed by population genetics modelling from the geographic distribution of disease variants. We first summarize some general approaches to identify disease alleles and the evolutionary characteristics of the variants, and we then introduce the recent genetic studies of IBD.

2.1 Genetic and evolutionary characteristics of diseases variants

Genetic mapping provides a powerful approach to identify genes and the biological processes underlying human diseases. From a vast amount of human genetic variations, a number of efforts have been made to identify those responsible for Mendelian diseases through linkage analysis. This method looks for a correlation between segregation (i.e. linkage) and phenotype within family pedigrees by using DNA sequence variants as genetic markers, without the need for prior an assumption about biological function. Since the causative variants for Mendelian diseases show high penetrance (proportion of individuals with a particular genotype who manifest a given phenotype), they tend to be transmitted into familial members in the Mendelian fashion. Thus, Mendelian diseases often feature a low frequency of causative mutations ($< 1\%$: rare variants) (Fig. 1a).

In contrast to Mendelian diseases, much less success has been achieved using linkage-based approaches for complex diseases (Risch and Merikangas 1996; Risch 2000; Altmuller et al. 2001). A single copy of the risk allele is sufficient to cause a Mendelian disease in the autosomal dominant pedigree, whereas the risk allele is neither sufficient nor necessary for complex diseases to occur (Knight 2009). The lack of Mendelian segregation of a complex disease in most families argues against the sufficiency of a mutation in any one gene (Chakravarti 1999). As alternative approaches to tracing transmission in families, one might localize disease's genes through association studies that compare the frequencies of genetic variants among affected and unaffected individuals. Based on a set of single nucleotide polymorphisms (SNPs) – which usually consists of several hundred thousand markers – a number of efforts have been made in GWAS. These scans show that the susceptibility variants involved in complex diseases have low or medium penetrance (i.e. incomplete penetrance) so that the disease variants are carried by healthy individuals, as well as by patients (Fig. 1a).

Based on these genetic properties of alleles, we can address the origins of disease variants in terms of evolutionary genetics. Mendelian diseases – which are usually due to highly penetrant and deleterious alleles that segregate in specific families – fit relatively simple *equilibrium* models of mutation-selection balance, in which disease alleles are removed by purifying selection and continually appearing through the mutation process (Di Rienzo 2006). These disease mutations are likely to have occurred relatively recently in human

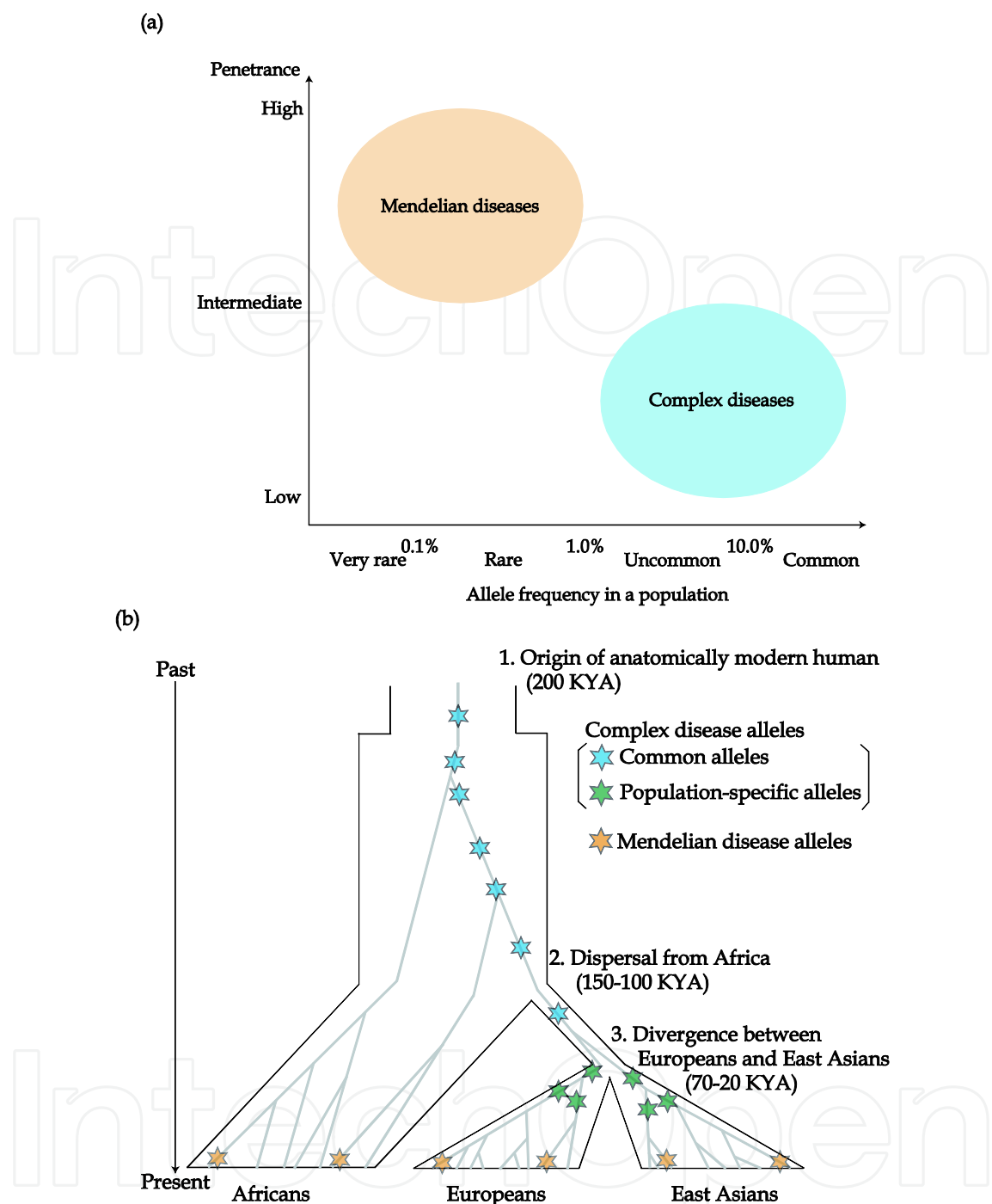


Fig. 1. The genetic properties of Mendelian disease and complex disease alleles. (a) Mendelian disease alleles have a high penetrance and their frequencies are very rare. In contrast, complex disease alleles are present both in patients and healthy controls in a population. This figure is modified from Box 7 in McCarthy et al. (2008). (b) The schematic diagram of human evolutionary history shows that anatomically modern humans originated in Africa around 200 thousand years ago (KYA). A small subset of ancestral populations dispersed from Africa around 150 to 100 KYA and then separated into Europeans and East Asians around 70 to 20 KYA. Mendelian disease alleles are likely to be recent mutations, while complex disease alleles – including common alleles and population-specific alleles – appeared before human population divergences

history (Fig. 1b). However, this *equilibrium* model is not applied to complex diseases, because of incomplete penetrance, gene-to-environmental interactions and polygenic inheritance. Most of the complex disease alleles spread across human populations, and these variants appeared before the divergence of populations. The complex disease alleles identified from GWAS are most common among geographically separated populations (e.g. Africans, Europeans and East Asians) (light-blue stars in Fig. 1b). Another phase of GWAS would be needed to focus on population-specific alleles (green stars in Fig. 1b) because of missing heritability (i.e. many of the genetic factors thought to be responsible for complex diseases can only explain between 5% and 50% of the diseases' heritability) and population-specific susceptibility. Specifically, these alleles have been jointly exposed to selective pressures and human demography, both of which are specific events to a particular geographic population. We describe population differences in susceptibility to IBD between Europeans and Japanese, and discuss evolutionary insights into the susceptibility to IBD.

2.2 Genetic studies of IBD in “European-ancestry populations”

Several studies have shown that an individual with IBD is more likely to have a relative with the disease (Budarf et al. 2009). Population-based studies find that 5-10% of patients have a first-degree family member with IBD, with the calculated sibling recurrence risk (λ_s : the ratio of the risk for the siblings of a patient to develop the disease compared to the risk for a general member of the same population) estimated to be 30-40 fold for CD and 10-20 fold for UC. The concomitant rate is significantly greater for monozygotic individuals than for dizygotic twins, for both CD (50-58% versus 0-12%) and UC (6-14% versus 0-5%) (Binder 1998). Hence, there are strong genetic contributions towards the risk of IBD, and especially for CD.

2.2.1 Genetic studies of CD

Generally, most of family-based studies have had limited success in finding genes for complex diseases, because of the non-Mendelian inheritance for the disease phenotypes. Nevertheless, linkage and positional cloning approaches have identified a nucleotide-binding oligomerization domain containing 2 (*NOD2*, also designated *CARD15* and *IBD1*) as the first susceptible gene for CD (Hugot et al. 2001; Ogura et al. 2001a). Moreover, the *IBD5* risk haplotype was identified from the linkage disequilibrium (LD) mapping of trios, and the risk haplotype included functionally interesting candidate genes: prolyl 4-hydroxylase (*P4HA2*), the interferon regulatory factor 1 (*IRF1*), and the organic cation transporter (OCTN) gene cluster (*SLC22A4* and *SLC22A5*, encoding OCTN1 and OCTN2, respectively) (Rioux et al. 2001; Peltekova et al. 2004). Since then, GWAS have been substantially improving our understanding of the biological pathways underlying the pathogenesis of CD, since the genetic contribution to CD is greater than to UC. A recent meta-analysis of the pooled data for six independent GWAS comprising 6,333 individuals with CD and 15,056 controls has reported 71 susceptibility loci (30 new susceptibility loci) to CD (Franke et al. 2010), and replicated the previously validated associations, including *NOD2* and *IBD5*, as well as other CD genes identified from several independent GWAS (Yamazaki et al. 2005; Duerr et al. 2006; Hampe et al. 2007; Libioulle et al. 2007; Parkes et al. 2007; Raelson et al. 2007; Rioux et al. 2007; The Wellcom Trust Case Control Consortium 2007; Barrett et al. 2008; Kugathasan et al. 2008). Therefore, the advent of GWAS has dramatically increased the number of susceptibility loci to CD in people of European descent.

2.2.2 Genetic studies of UC

CD and UC share many diagnostic features, and relatives with CD or UC are likely to be at an increased risk of developing either form of IBD, indicating the existence of both phenotype-specific and shared susceptibility loci for CD and UC. Before GWAS in UC, candidate gene approaches to the susceptibility loci identified in CD were conducted in UC. It was shown that several genes – including the interleukin-23 receptor (*IL23R*), the NK2 transcription-factor-related, locus 3 (*NKX2-3*) and the macrophage stimulating 1 (*MST1*) – were also significantly associated with UC, whereas the other representative genes in CD, such as *NOD2*, showed no susceptibility to UC (Fisher et al. 2008; Franke et al. 2008). Subsequently, GWAS – as well as these candidate gene association studies – have identified 18 susceptibility loci for UC, including established risk loci specific to UC, such as the hepatocyte nuclear factor 4 α (*HNF4A*), the cadherin 1 (*CDH1*) and the laminin β 1 subunit (*LAMB1*) that are highlighted with the role of defective barrier function in UC pathogenesis (Barrett et al. 2009). A meta-analysis of six GWAS datasets of UC, comprising 6,687 cases and 19,718 controls, has identified 29 additional risk loci, increasing the number of UC loci to 47 (Anderson et al. 2011). The total number of confirmed IBD risk loci is about 99, and a minimum of 28 show shared association signals between CD and UC. Thus, recent GWAS successes have accelerated our knowledge about the commonalities and unique features of the aetiology between CD and UC.

2.3 Genetic studies of IBD in Japanese (“non-European-ancestry”) populations

Many GWAS’s efforts to identify susceptibility genes of IBD have been successful in European-ancestry populations, and recent advances have provided substantial insights into the maintenance of mucosal immunity and the pathogenesis of IBD (Xavier and Podolsky 2007). Furthermore, some of the IBD variants have also had their pathogenic roles demonstrated by *in vivo* and *in vitro* functional studies. *NOD2*, one of the established susceptibility genes to CD (Hugot et al. 2001; Ogura et al. 2001b), encodes a protein that recognizes pathogen-associated molecular patterns: common motifs of the peptidoglycan product muramyl dipeptide (MDP), which modulates both innate and adaptive immune responses (Shaw et al. 2011). The cytosine insertion in *NOD2* exon 11 (3,020C) results in a frameshift and generates a truncated *NOD2* protein (1,007 of 1,040 amino acid residues), which induces impaired activation of the transcriptional factor NF- κ B (Chamaillard et al. 2003). For *ATG16L* encoding a key autophagy molecule, the patients with a homozygote of alanine at 300 amino acid residues (T300A) display disorganised or diminished granules in Paneth cells, which are specialised epithelial cells for controlling the intestinal environment by the release of granules (Cadwell et al. 2008). Several additional lines of evidence, including transgenic mouse experiments for the susceptibility genes (Cadwell et al. 2008; Cadwell et al. 2010; Travassos et al. 2010), show a functional deficiency of CD variants. However, significant associations between these genomic regions and IBD have not been detected in non-European populations, suggesting that there is a population-specific susceptibility to IBD.

The incidence of IBD has been increasing substantially within the Japanese population. The prevalence rate has risen seven times between 1,985 and 2,006 (Hilmi et al. 2006). The disease-mapping strategy of early association studies in Japan was to test the susceptibilities of genes reported from European-ancestry populations using Japanese IBD patients and controls. The candidate gene study shows that the *NOD2* variants statistically and functionally confirmed in Europeans are completely absent – both in patients and controls – in the Japanese population (Yamazaki et al. 2002). This result is consistent with the other

East Asian populations, such as Korean (Croucher et al. 2003) and Chinese (Guo et al. 2004). To elucidate the similarities and differences of susceptibility to IBD between Europeans and the Japanese, associations with CD and UC have been tested for seven susceptible genomic regions, including *NOD2*, *IL23R*, *ATG16L1*, *TNFSF15*, *SLC22A4*, *IRGM*, and 10q21 (Nakagome et al. 2010). Each of these genomic regions, which have been confirmed by multiple independent GWAS and the meta-analysis described above, is known to be associated with CD in European-ancestry populations. Moreover, Nakagome et al. (2010) have focused on a local differences in susceptibility to IBD within Japanese populations, because the population stratification (differences in allele frequencies between sub-populations due to different ancestry) is observed between Honshu (the eastern area of Japan's main-island) and Kyushu-Okinawa (the southwestern islands of the Japanese archipelago) (Yamaguchi-Kabata et al. 2008). Afterwards, the association of nine SNPs located in the seven genomic regions was examined in the Kyushu population, consisting of 130 individuals with CD, 82 individuals with UC, and 168 controls (Table 1), and which was also compared with the genotype data from the European and Honshu Japanese populations.

2.3.1 Differences in susceptibility to IBD between Europeans and Japanese

The samples acquired from each of the Kyushu Japanese subjects are first analysed to determine the genotypes of the nine SNPs, which are previously identified from the CD-associated genomic regions (Table 1). These samples are also examined in order to determine any association of CD or UC with the risk alleles that had been identified in European-ancestry populations. Table 2 illustrates the genotype frequencies as well as allele frequencies between cases and controls. The p -values of the χ^2 test corrected by the permutation test are shown in Table 3. In the Kyushu Japanese subjects, the risk alleles for *NOD2* (rs2066844, rs2066845, rs2066847) and *SLC22A4* (rs1050152) are not found to be present. The analysis of the *IL23R* (rs11209026) gene shows that the risk allele is fixed in the CD cases, the UC cases and the control. Thus, these SNPs are not polymorphic in the Japanese population, which confirms previous studies (Yamazaki et al. 2002; Yamazaki et al. 2004; Yamazaki et al. 2007). In contrast, the remaining four SNP sites (i.e. rs2241880, rs3810936, rs10065172, and rs10761659) are shown to be polymorphic (Table 2). Furthermore, three of the SNPs, including rs2241880 in *ATG16L1*, rs10065172 in *IRGM*, and rs10761659 in 10q21, do not show any significant association with either CD or UC (Table 3). Only one SNP site was found to be significantly associated with CD (p -value = 0.047) and UC (p -value = 0.050) is in the *TNFSF15* gene (rs3810936). The odds ratio (OR) of the risk allele ("G" allele in rs3810936) is 1.551 (95% CI: 1.090-2.207) for CD and 1.692 (95% CI: 1.117-2.562) for UC (Table 3).

The differences in allele frequency between the Kyushu and Honshu subjects are further compared by an χ^2 test that included the Kyushu (K-) CD, the Honshu (H-) CD and each control (2-by-2 pairs: K-CD and H-CD, K-controls and H-controls, K-CD and H-controls, K-controls and H-CD) (Fig. 2). A significant association between *TNFSF15* (rs3810936) and CD is detected in all pairs of CD-controls from both the Kyushu and Honshu subjects. Previous studies of Honshu Japanese subjects have examined a greater number of SNPs in the other six genome regions, compared to Nakagome et al. (2010), but do not detect a significant association between CD and these SNPs (Yamazaki et al. 2002; Yamazaki et al. 2004; Yamazaki et al. 2007; Yamazaki et al. 2009). As an exception, Yamazaki et al. (2005) identified *TNFSF15* as a CD susceptibility gene in the Honshu subjects, due to a significant association ($p < 0.0001$) between 20 SNP sites and CD. Using the Kyushu subjects, Nakagome et al. (2010) supports the previous results which demonstrated that the CD

| Susceptibility genes | No. of previous studies ^a | SNP IDs genotyped in this study | SNP type | Amino acid changes |
|----------------------|--------------------------------------|---------------------------------|----------------------|--------------------|
| NOD2 | 1, 2, 5, 6, 9, 10, 11, and 12 | rs2066844 | Nonsynonymous | Arg > Trp |
| | | rs2066845 | Nonsynonymous | Gly > Arg |
| | | rs2066847 | Insertion (C allele) | frame shift |
| IL23R | 5, 7, 9, 10, 11, and 12 | rs11209026 | Nonsynonymous | Ala > Gln |
| ATG16L1 | 6, 9, 11, and 12 | rs2241880 | Nonsynonymous | Thr > Ala |
| TNFSF15 | 4, 11, and 12 | rs3810936 | Synonymous | - |
| SLC22A4 | 3, 6, 8, 10, and 11 | rs1050152 | Nonsynonymous | Leu > Phe |
| IRGM | 8 and 11 | rs10065172 | Synonymous | - |
| 10q21 | 9 and 11 | rs10761659 | Non-coding | - |

^a The numbers are corresponding to previous GWAS: 1. Hugot et al. (2001); 2. Ogura et al. (2001); 3. Peltekova et al. (2004); 4. Yamazaki et al. (2005); 5. Duerr et al. (2006); 6. Hampe et al. (2007); 7. Libioulle et al. (2007); 8. Parkes et al. (2007); 9. Raelson et al. (2007); 10. Rioux et al. (2007); 11. The Wellcom Trust Case Control Consortium (2007); 12. Kugathasan et al. (2008).

Table 1. Susceptibility genes of IBD

Table 2. Number of genotypes and alleles in the Kyushu subjects

| Susceptibility genes | SNP IDs | Alleles | Risk allele A ^a | Genotype frequency ^b | | | | | | | | |
|----------------------|------------|---------|----------------------------|---------------------------------|----|-----|-----|----|-----|----|----------|----|
| | | | | CD | | | | | | | Controls | |
| | | | | AA | Aa | aa | UC | AA | Aa | aa | AA | Aa |
| NOD2 | rs2066844 | C/T | T | 0 | 0 | 100 | 0 | 0 | 100 | 0 | 0 | 0 |
| | rs2066845 | G/C | C | 0 | 0 | 100 | 0 | 0 | 100 | 0 | 0 | 0 |
| | rs2066847 | -/C | C insertion | 0 | 0 | 100 | 0 | 0 | 100 | 0 | 0 | 0 |
| IL23R | rs11209026 | A/G | G | 100 | 0 | 0 | 100 | 0 | 0 | 0 | 100 | 0 |
| ATG16L1 | rs2241880 | A/G | G | 8 | 40 | 53 | 7 | 32 | 61 | 7 | 39 | 39 |
| TNFSF15 | rs3810936 | A/G | G | 49 | 47 | 4 | 57 | 34 | 9 | 39 | 49 | 49 |
| SLC22A4 | rs1050152 | C/T | T | 0 | 0 | 100 | 0 | 0 | 100 | 0 | 0 | 0 |
| IRGM | rs10065172 | T/C | C | 36 | 47 | 16 | 40 | 44 | 16 | 37 | 51 | 51 |
| 10q21 | rs10761659 | A/G | G | 54 | 36 | 10 | 46 | 44 | 10 | 47 | 42 | 42 |

^a Associated alleles were referred from the previous studies and annotated as "A " .
^b The genotypes were shown as "AA " homozygote of risk alleles, "Aa " heterozygote of risk and non-risk alleles, and "aa " homozygote of non-risk alleles.

susceptibility genes identified in Europeans are not significantly associated with CD in the Japanese population. The data on the *TNFSF15* gene using the Kyushu subjects also indicated a significant positive association with CD. Therefore, these results strongly support *TNFSF15* as a CD susceptibility gene in the Japanese population.

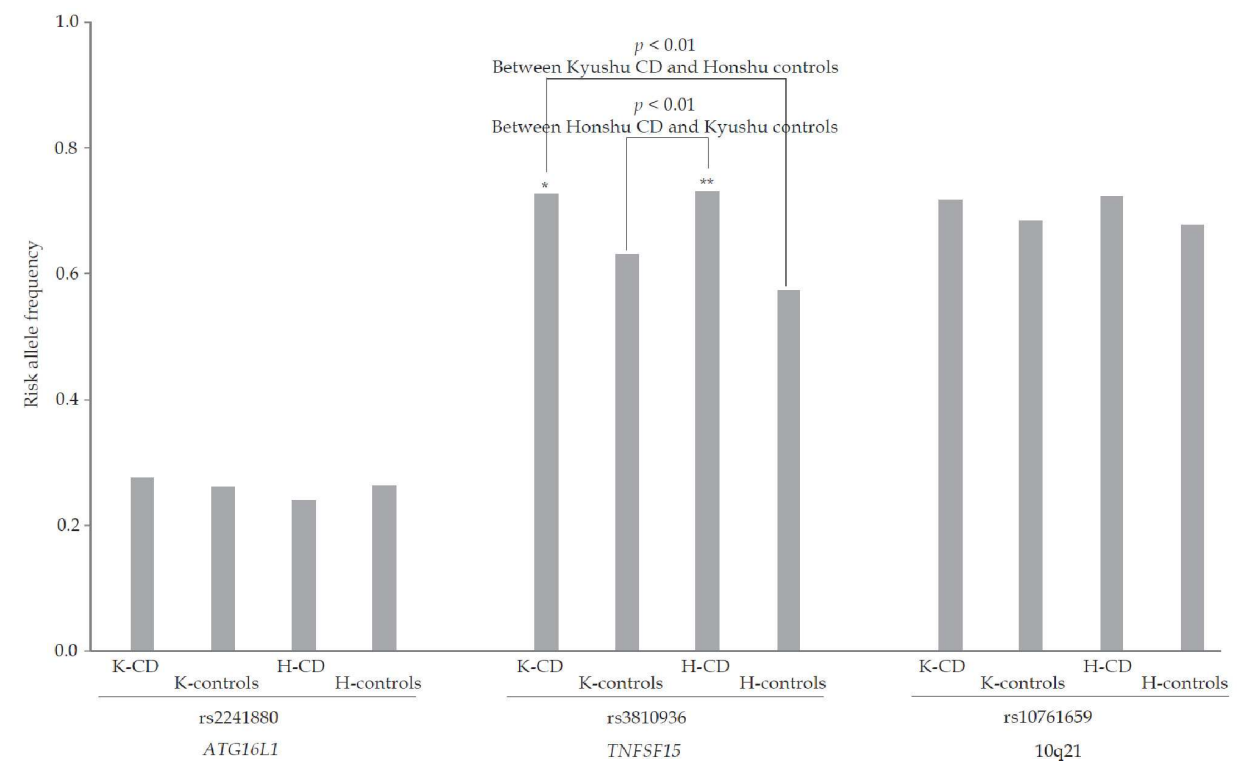


Fig. 2. The distributions of risk allele frequency for three SNPs (rs2241880, rs3810936, rs10761659) in the Kyushu and Honshu subjects. The abbreviations, K-CD, K-controls, H-CD and H-controls indicate Kyushu CD, Kyushu-controls, Honshu CD, and Honshu-controls. The asterisks above the columns indicate a significant allele frequency difference between cases and controls in the same region (p -values for the Kyushu subject are referred from Table 3 and those for the Honshu subject are referred from Yamazaki et al. (2005)). A single asterisk denotes $p < 0.05$, and double asterisks denote $p < 0.01$

2.3.2 Genotype association with CD and UC

Disease alleles on autosomes are present as heterozygote or homozygote in an individual, and their effects can be categorised as dominant, recessive or additive. The genotype association of the four polymorphic SNP sites is tested in CD or UC, respectively. No significant association of *ATG16L1* (rs2241880), *IRGM* (rs10065172), or 10q21 (rs10761659) is observed with either CD or UC. The different significances of *TNFSF15* (rs3810936) between CD and UC are shown in the recessive model (i.e. AA vs. Aa + aa) and the dominant model (i.e. AA + Aa vs. aa) (Table 3). In the recessive model, the test shows a significant association between the homozygote for the risk allele and UC (p -value = 0.019). The OR of the risk allele homozygote is determined to be 2.132 (95% CI: 1.247 – 3.644). In the dominant model, both the heterozygote and the homozygote for the risk allele are found to be significantly associated with CD (p -value = 0.025 and OR = 3.497, 95% CI: 1.341 – 9.117). Thus, the effect of susceptibility in the *TNSF15* risk allele (rs3810936) is likely to be different between CD and UC of the Kyushu Japanese.

Table 3. Associations of alleles and genotypes with Crohn’s disease or ulcerative colitis in the Kyushu subjects

| Susceptibility genes | SNP IDs | Allele (<i>A</i> / <i>a</i>) | | Genotype (<i>AA</i> / <i>Aa</i> / <i>aa</i>) | | | | | | | | Disease |
|----------------------|------------|--------------------------------|--------|--|--------------------------|------------------|--------|--|--------------------------|-----------------|--|---------|
| | | | | Dominant model (<i>A</i> vs <i>a</i>) | | | | Recessive model (<i>AA</i> vs others) | | | | |
| | | <i>p</i> -values | | Odds ratio (95% CI) | | <i>p</i> -values | | Odds ratio (95% CI) | | <i>p</i> -value | | |
| | | CD | UC | CD | UC | CD | UC | CD | UC | CD | | |
| <i>ATG16L1</i> | rs2241880 | 0.986 | 0.948 | 1.077 (0.744 - 1.557) | 0.855 (0.551 - 1.327) | 0.997 | 0.767 | 1.161 (0.477 - 2.828) | 1.091 (0.388 - 3.068) | 1.000 | | |
| <i>TNFSF15</i> | rs3810936 | 0.047* | 0.050* | 1.551 (1.090 - 2.207) | 1.692 (1.117 - 2.562) | 0.232 | 0.019* | 1.539 (0.966 - 2.454) | 2.132 (1.247 - 3.644) | 0.001 | | |
| <i>IRGM</i> | rs10065172 | 0.921 | 1.000 | 0.886 (0.634 - 1.237) | 0.987 (0.670 - 1.454) | 0.999 | 0.974 | 0.956 (0.592 - 1.543) | 1.137 (0.660 - 1.960) | 0.999 | | |
| 10q21 | rs10761659 | 0.849 | 1.000 | 1.171 (0.818 - 1.674) | 0.995 (0.664 - 1.490) | 0.741 | 0.999 | 1.284 (0.808 - 2.041) | 0.965 (0.566 - 1.644) | 1.000 | | |

* *p* -values < 0.05.

2.3.3 Differences in the Genotype Relative Risk between CD and UC

The genotype relative risk (GRR) test (Schaid and Sommer 1993) is conducted for the SNP site of rs3810936 (*TNFSF15*) that is found to be significantly associated with CD or UC. This test is adopted so as to determine whether the dominant, recessive or additive model best fit the observed genotype frequency using the likelihood ratio test. Based on the likelihood method, the GRR test is modified to include unrelated individuals and also low penetrance variants.

Let A be the disease-associated allele, and let a be the non disease-associated allele. The probability of disease is defined, conditional on the genotype at the particular SNP site, as:

$$f_{2case} = P(D | AA), f_{1case} = P(D | Aa), f_{0case} = P(D | aa), \quad (1)$$

where D is the event that an individual has the disease. The complex disease allele is retained not only in cases but also in healthy controls. Next, the probability of non-disease is defined, conditional on the genotype at the particular SNP site, as:

$$f_{icontrol} = 1 - f_{icase} (i = 0, 1, 2), \quad (2)$$

Assuming that a population is in the Hardy-Weinberg equilibrium, then by Bayes' rule we have:

$$P(AA | D) = \frac{\psi_{2case} p^2}{R_{case}}, P(Aa | D) = \frac{\psi_{1case} 2pq}{R_{case}}, P(aa | D) = \frac{q^2}{R_{case}}, \quad (3)$$

$$P(AA | N) = \frac{\psi_{2control} p^2}{R_{control}}, P(Aa | N) = \frac{\psi_{1control} 2pq}{R_{control}}, P(aa | N) = \frac{q^2}{R_{control}}, \quad (4)$$

where,

$$\psi_{2case} = \frac{f_{2case}}{f_{0case}}, \psi_{1case} = \frac{f_{1case}}{f_{0case}}, R_{case} = \psi_{2case} p^2 + \psi_{1case} 2pq + q^2, \quad (5)$$

$$\psi_{2control} = \frac{f_{2control}}{f_{0control}}, \psi_{1control} = \frac{f_{1control}}{f_{0control}}, R_{control} = \psi_{2control} p^2 + \psi_{1control} 2pq + q^2, \quad (6)$$

p is the population frequency of A , q is the frequency of a , and N is the non-disease state. Note that the likelihood depends on four independent parameters, ψ_{2case} , ψ_{1case} , p_{case} and f_{0case} . A standard numerical maximisation procedure can then be used to find the maximum-likelihood estimates of ψ_{2case} , ψ_{1case} , p_{case} and f_{0case} with the likelihood

$$L(\psi_{2case}, \psi_{1case}, p, f_{0case}) = P(AA | D)^{n_{AAcase}} P(Aa | D)^{n_{Aacase}} P(aa | D)^{n_{aacase}} \\ P(AA | N)^{n_{AAcontrol}} P(Aa | N)^{n_{Aacontrol}} P(aa | N)^{n_{aacontrol}}, \quad (7)$$

where n_{AAcase} , n_{Aacase} , n_{aacase} or $n_{AAcontrol}$, $n_{Aacontrol}$, $n_{aacontrol}$ are the observed numbers of cases or controls exhibiting each genotype.

We considered the following four hypotheses:

- i. No association of the marker with disease, H_0 : $\psi_{1case} = \psi_{2case} = 1$,
- ii. Dominant disease expression, H_D : $\psi_{1case} = \psi_{2case} = \psi$,
- iii. Recessive disease expression, H_R : $\psi_{1case} = 1$,
- iv. Additive disease expression, H_A : $\psi_{1case} = \psi$, $\psi_{2case} = 2\psi$.

The maximum likelihood estimate can be found by maximising the likelihood function under each hypothesis with the condition: $0 < p < 1$, $1 < \psi_{1case}$, $1 < \psi_{2case}$, and $0 < f_{0case} < 1$. As described by Scaid and Sommer (1993), we next adopted the likelihood-ratio test (LRT). The LRT statistics for testing the hypothesis of no association is:

$$\chi^2_2 = 2\log[L(\hat{p}, \hat{\psi}_{2case}, \hat{\psi}_{1case}, \hat{f}_{0case})/L(p, 1, 1, f_{0case})], \quad (8)$$

The statistics for testing the hypothesis of a dominant model are:

$$\chi^2_1 = 2\log[L(\hat{p}, \hat{\psi}_{2case}, \hat{\psi}_{1case}, \hat{f}_{0case})/L(p, \psi, \psi, f_{0case})], \quad (9)$$

The statistics for testing the hypothesis of a recessive model are:

$$\chi^2_1 = 2\log[L(\hat{p}, \hat{\psi}_{2case}, \hat{\psi}_{1case}, \hat{f}_{0case})/L(p, 1, \psi_{2case}, f_{0case})], \quad (10)$$

Finally, the statistics for testing the hypothesis of an additive model are:

$$\chi^2_1 = 2\log[L(\hat{p}, \hat{\psi}_{2case}, \hat{\psi}_{1case}, \hat{f}_{0case})/L(p, \psi, 2\psi, f_{0case})]. \quad (11)$$

For the SNP site (rs3810936) in *TNFSF15*, four parameters – including ψ_{2case} , ψ_{1case} , p_{case} , and f_{0case} – are inferred from the general likelihood equation so that the likelihood ratio test can be applied to (7). Based on these hypotheses, the CD likelihood ratio test rejects the null (LRT p -value = 0.005), recessive (LRT p -value = 0.024) and additive models (LRT p -value = 0.021) (Table 4). However, the dominant model was not rejected (LRT p -value = 0.361). In contrast, the UC likelihood ratio test was not found to reject any of the models, most likely due to a lack of power. Nevertheless, the recessive model demonstrates a better fit with the observed genotype frequency than do the dominant and additive models (LRT p -values are not assessed because the same LogL value is in both the recessive and the full model). These results are also supported by Akaike information criterion (AIC) values, which show the minimum values of the CD dominant model (AIC = 540.562) and of the UC recessive model (AIC = 470.516) (Table 4). Thus, the statistical test for the genotype risk of rs3810936-G showed that the CD mGRR data best fits the dominant model, while the UC mGRR data best fits the recessive model.

The similarities and differences between CD and UC are thought to be important in understanding the pathogenesis of each disease (Dubois and van Heel 2008). The results from the mGRR tests clearly show that the genotype of the rs3810936-G allele in *TNFSF15* exhibits a different effect in CD compared to UC. The risk of the rs3810936-G allele was determined to be comparable between CD and UC (OR: 1.551 fold in CD and 1.692 fold in UC) (Table 3). However, the genotype relative risk between CD and UC was found to be greatly different (the risk of “GA” or “GG”: 3.604 fold or 4.310 fold in C; no risk or 1.679 fold in UC) (Table 4). Thus, it is likely that the risk variant of *TNFSF15* functions as a “dominant” allele in CD, whereas it functions as a “recessive” allele in UC.

Table 4. Modified genotype relative risk (mGRR) test of rs3810936 for CD and UC in the Kyushu subjects

| Model | Crohn's disease | | | | | | | Ulcerative colitis | | |
|-----------------------|-----------------|---------------------------|---------------------------|----------|---------------------------|------------------|----------------------|--------------------|---------------------------|---------------------------|
| | Log <i>L</i> | <i>ψ</i> _{1case} | <i>ψ</i> _{2case} | <i>p</i> | <i>f</i> _{0case} | AIC ^a | LRT <i>p</i> -values | Log <i>L</i> | <i>ψ</i> _{1case} | <i>ψ</i> _{2case} |
| Full | -266.863 | 3.604 | 4.310 | 0.635 | 0.006 | 541.726 | - | -232.258 | 1.000 | 1.679 |
| Null (no association) | -272.188 | - | - | 0.674 | 0.881 | 548.376 | 0.005 ^{**} | -236.100 | - | - |
| Dominant | -267.281 | 3.580 | 3.580 | 0.646 | 0.000 | 540.562 | 0.361 | -235.704 | 1.310 | 1.310 |
| Recessive | -269.425 | - | 1.000 | 0.714 | 1.000 | 544.850 | 0.024 [*] | -232.258 | - | 1.679 |
| Additive | -269.512 | 3.383 | 6.766 | 0.598 | 0.000 | 545.024 | 0.021 [*] | -232.390 | 1.000 | 2.000 |

^a The AIC is -2log *L* + 2*k* , where *k* is the number of estimated parameters.
^b *N. A.* means "not assessed" (there is no LRT *p*-values, because Log*L* value in Recessive model is equal to that in Full model).
^{*} *p* -values < 0.05
^{**} *p* -values < 0.01

2.3.4 Population-specific susceptibility to IBD

Given that CD and UC inheritance do not follow the typical Mendelian fashion, it is possible that the dominant/recessive effect of the risk allele (i.e. the biological roles of the disease etiology), as suggested by the mGRR test above, is shaped by the multiple genetic and environmental factors that are involved in each disease's mechanism. Recently, it has been argued that the reciprocal interaction between multiple genetic and environmental factors is important for complex diseases (Emison et al. 2005). The CD risk alleles identified in Europeans are absent – or not associated with CD – in the Japanese population, except for *TNFSF15*. Ethnic group-specific susceptibility is not a novel idea, and has generally been observed in complex diseases (Altshuler et al. 2008; Rosenberg et al. 2010; Bustamante et al. 2011). Furthermore, the CD risk allele in *TNFSF15* is also associated with UC in the Kyushu population. However, a previous study has found no association with UC at the same SNP in *TNFSF15*, using non-Kyushu Japanese subjects (Kakuta et al. 2006). These results imply that UC, as well as CD, may exhibit a population-specific susceptibility within the various Japanese subjects, similar to the susceptibility shown to exist between Japanese and European subjects. Using world-wide population samples, Myles et al. (2008) have examined the allele frequencies of complex diseases, including CD, and indicated the importance of geographic variation in disease-associated alleles (Myles et al. 2008). Specifically, allele frequencies have commonly been observed to be gradients among human populations (Tishkoff et al. 1996; Oota et al. 2004; Tishkoff and Kidd 2004; The International HapMap Consortium 2005; The International HapMap Consortium 2007), including the Japanese population (Yamaguchi-Kabata et al. 2008). Hence, allele frequencies fluctuate in each geographic region where environmental factors are different, and the population-specific susceptibility of complex diseases might be reflected by subdivisions (i.e. subpopulations) among human populations. Population-specific genetic and environmental factors are likely to cause dominant or recessive genotype risks in each subpopulation. Therefore, it is necessary to investigate a greater number of local cases and controls in order to reveal the mechanisms of complex diseases.

2.4 Evolutionary Insights into the geographic variation of IBD risk alleles

Most of the genetic variants susceptible to complex diseases, including IBD, appeared before the divergences of human populations (Fig. 1b), and geographic gradients in allele frequencies are attributed to differences in evolutionary history, such as migration, changes in population size and natural selection among the populations. A representative example of a disease-causative allele maintained by environmental adaptation is that associated with sickle-cell anaemia, which causes severe anaemia, but its frequency is highly maintained in particular geographical regions because it confers resistance to malaria (Pasvol et al. 1978). Another example is the thrifty genotype hypothesis (Neel 1962) that the genetic predisposition to type II diabetes is the consequence of metabolic adaptations to an ancient lifestyle characterised by a fluctuating and unpredictable food supply and high levels of physical activity. With the switch to a sedentary lifestyle and energy-dense diets in civilised countries, the thrifty genotype is no longer advantageous and gives rise to disease phenotypes. The genetic variants that have been found to confer IBD risks point to the importance of innate immunity, autophagy and phagocytosis in IBD pathogenesis. Hence, we attribute population-specific IBD susceptibility to natural selection on the IBD variants,

or the other variation(s) closely linked to the IBD risk alleles (i.e. genetic hitchhiking) which might be an adaptation to pathogen infections in ancestral human populations. Indeed, CD risk alleles in *NOD2* (rs2066844, rs2066845, and rs2066847) and in *SLC22A4* (rs1050152) – all of them involve amino acid changes or frameshift – are completely absent in Japanese populations (Table 2), and possibly specific to European populations. These alleles might have been maintained only in European-ancestry populations by natural selection. To reveal the mechanisms of how IBD risk alleles are spread throughout human populations, in the near future we would like to collect detailed polymorphism data from geographically diverse populations so as to conduct evolutionary and population genetics analysis to detect the signal(s) of natural selection.

Evolutionary insights into the genetic properties of individual risk alleles are useful for understanding how we can translate the results from GWAS in one population to other diverse populations so as to understand the similarities and differences among GWAS results in geographically separated populations. Hence, population genetic data and modelling efforts have had important roles in the characterization of disease alleles and in the planning of GWAS, respectively. More importantly, the current disease allele frequency and heterogeneity in IBD susceptible alleles among populations are the products of a long-term human evolutionary history, and population-specific variants are likely to confer substantial risks of IBD in a particular population. Therefore, population-genetic modelling of the IBD risk alleles will highlight quantitative differences in population-specific mutations and provide a comprehensive catalogue of the intermediate variants, which are neither rare (< 1% in a population) nor common (> 5% in any populations), specific to a geographic region (Fig. 1b).

3. Conclusion

Among complex diseases, GWAS have produced numerous successes in identifying genes and genetic loci that contribute to IBD susceptibility. Despite distinct clinical features, approximately 30% of IBD-associated loci (28/99) are shared between UC and CD, indicating that these diseases engage common pathways and may be part of a mechanistic continuum (Khor et al. 2011). However, findings from GWAS in European-ancestry populations are not always easily translated into the rest of the world, which implies that the IBD-relevant biological pathways are different among geographic populations. Medical and evolutionary approaches are essential in the next phase of researches on human complex diseases. As to the medical studies, GWAS and whole-genome sequence approaches should incorporate geographically diverse populations and provide a comprehensive catalogue of candidate risk alleles. The important idea is that the worldwide human population and its distribution of disease-risk variation represent a singular outcome of human evolutionary history which will underlie future disease-mapping studies. Population-genetic modelling for disease alleles has become important in unravelling the reasons why the risk of IBD has remained within human populations. Evolutionary research should translate the outputs from large-scale medical studies to biological interpretations for the genetic backgrounds of IBD. As the technological barriers to the production of genomic data continue to fall, it can be hoped that the human genomics community will accept the challenge of capitalising on the full range of human diversity in the next wave of investigations of the variants that underlie IBD, since the expansion of our

understanding for human diversity is significant in the examination of any new aspects of genetic variation associated with IBD's pathogenesis. Therefore, evolutionary insights into IBD genetics will give a way to a paradigm of understanding inter-population and inter-individual differences in these diseases mechanisms and of developing personalised medicine for the prevention of and care for IBD sufferers.

4. Acknowledgement

This was supported by a Grant-in-Aid for Scientific Research (C) from JSPS (19570226) to HO, by a Grant-in-Aid for Scientific Research (B) from JSPS (21370108) to HO, and by a Grant-in-Aid for JSPS Research fellow (21-7453) to SN.

5. References

- Altmuller, J., L. J. Palmer, G. Fischer, H. Scherb, and M. Wjst. 2001. Genomewide scans of complex human diseases: true linkage is hard to find. *Am J Hum Genet* 69:936-950.
- Altshuler, D., M. J. Daly, and E. S. Lander. 2008. Genetic mapping in human disease. *Science* 322:881-888.
- Anderson, C. A.G. Boucher C. W. Lees A. Franke M. D'Amato K. D. Taylor et al. 2011. Meta-analysis identifies 29 additional ulcerative colitis risk loci, increasing the number of confirmed associations to 47. *Nat Genet* 43:246-252.
- Barrett, J. C., S. Hansoul, D. L. Nicolae, J. H. Cho, R. H. Duerr, J. D. Rioux et al. 2008. Genome-wide association defines more than 30 distinct susceptibility loci for Crohn's disease. *Nat Genet* 40:955-962.
- Barrett, J. C., J. C. Lee, C. W. Lees, N. J. Prescott, C. A. Anderson, A. Phillips et al. 2009. Genome-wide association study of ulcerative colitis identifies three new susceptibility loci, including the HNF4A region. *Nat Genet* 41:1330-1334.
- Binder, V. 1998. Genetic epidemiology in inflammatory bowel disease. *Dig Dis* 16:351-355.
- Budarf, M. L., C. Labbe, G. David, and J. D. Rioux. 2009. GWA studies: rewriting the story of IBD. *Trends Genet* 25:137-146.
- Bustamante, C. D., E. G. Burchard, and F. M. De la Vega. 2011. Genomics for the world. *Nature* 475:163-165.
- Cadwell, K., J. Y. Liu, S. L. Brown, H. Miyoshi, J. Loh, J. K. Lennerz et al. 2008. A key role for autophagy and the autophagy gene Atg16l1 in mouse and human intestinal Paneth cells. *Nature* 456:259-263.
- Cadwell, K., K. K. Patel, N. S. Maloney, T. C. Liu, A. C. Ng, C. E. Storer, R. D. Head, R. Xavier, T. S. Stappenbeck, and H. W. Virgin. 2010. Virus-plus-susceptibility gene interaction determines Crohn's disease gene Atg16L1 phenotypes in intestine. *Cell* 141:1135-1145.
- Chakravarti, A. 1999. Population genetics--making sense out of sequence. *Nat Genet* 21:56-60.
- Chamaillard, M., D. Philpott, S. E. Girardin, H. Zouali, S. Lesage, F. Chareyre, T. H. Bui, M. Giovannini, U. Zaehring, V. Penard-Lacronique, P. J. Sansonetti, J. P. Hugot, and G. Thomas. 2003. Gene-environment interaction modulated by allelic heterogeneity in inflammatory diseases. *Proc Natl Acad Sci U S A* 100:3455-3460.

- Croucher, P. J., S. Mascheretti, J. Hampe, K. Huse, H. Frenzel, M. Stoll, T. Lu, S. Nikolaus, S. K. Yang, M. Krawczak, W. H. Kim, and S. Schreiber. 2003. Haplotype structure and association to Crohn's disease of CARD15 mutations in two ethnically divergent populations. *Eur J Hum Genet* 11:6-16.
- Di Rienzo, A. 2006. Population genetics models of common diseases. *Curr Opin Genet Dev* 16:630-636.
- Dubois, P. C., and D. A. van Heel. 2008. New susceptibility genes for ulcerative colitis. *Nat Genet* 40:686-688.
- Duerr, R. H., K. D. Taylor, S. R. Brant, J. D. Rioux, M. S. Silverberg, M. J. Daly et al. 2006. A genome-wide association study identifies IL23R as an inflammatory bowel disease gene. *Science* 314:1461-1463.
- Emison, E. S., A. S. McCallion, C. S. Kashuk, R. T. Bush, E. Grice, S. Lin, M. E. Portnoy, D. J. Cutler, E. D. Green, and A. Chakravarti. 2005. A common sex-dependent mutation in a RET enhancer underlies Hirschsprung disease risk. *Nature* 434:857-863.
- Fisher, S. A., M. Tremelling, C. A. Anderson, R. Gwilliam, S. Bumpstead, N. J. Prescott et al. 2008. Genetic determinants of ulcerative colitis include the ECM1 locus and five loci implicated in Crohn's disease. *Nat Genet* 40:710-712.
- Franke, A., T. Balschun, T. H. Karlsen, J. Hedderich, S. May, T. Lu, D. Schuldt, S. Nikolaus, P. Rosenstiel, M. Krawczak, and S. Schreiber. 2008. Replication of signals from recent studies of Crohn's disease identifies previously unknown disease loci for ulcerative colitis. *Nat Genet* 40:713-715.
- Franke, A., D. P. McGovern, J. C. Barrett, K. Wang, G. L. Radford-Smith, T. Ahmad et al. 2010. Genome-wide meta-analysis increases to 71 the number of confirmed Crohn's disease susceptibility loci. *Nat Genet* 42:1118-1125.
- Guo, Q. S., B. Xia, Y. Jiang, Y. Qu, and J. Li. 2004. NOD2 3020insC frameshift mutation is not associated with inflammatory bowel disease in Chinese patients of Han nationality. *World J Gastroenterol* 10:1069-1071.
- Hampe, J., A. Franke, P. Rosenstiel, A. Till, M. Teuber, K. Huse et al. 2007. A genome-wide association scan of nonsynonymous SNPs identifies a susceptibility variant for Crohn disease in ATG16L1. *Nat Genet* 39:207-211.
- Hilmi, I., Y. M. Tan, and K. L. Goh. 2006. Crohn's disease in adults: observations in a multiracial Asian population. *World J Gastroenterol* 12:1435-1438.
- Hugot, J. P., M. Chamaillard, H. Zouali, S. Lesage, J. P. Cezard, J. Belaiche et al. 2001. Association of NOD2 leucine-rich repeat variants with susceptibility to Crohn's disease. *Nature* 411:599-603.
- Kakuta, Y., Y. Kinouchi, K. Negoro, S. Takahashi, and T. Shimosegawa. 2006. Association study of TNFSF15 polymorphisms in Japanese patients with inflammatory bowel disease. *Gut* 55:1527-1528.
- Khor, B., A. Gardet, and R. J. Xavier. 2011. Genetics and pathogenesis of inflammatory bowel disease. *Nature* 474:307-317.
- Knight, C. J. 2009. *Human Genetic Diversity: Functional consequences for health and disease*. Oxford University Press Inc, New York, NY.
- Kugathasan, S., R. N. Baldassano, J. P. Bradfield, P. M. Sleiman, M. Imielinski, S. L. Guthery et al. 2008. Loci on 20q13 and 21q22 are associated with pediatric-onset inflammatory bowel disease. *Nat Genet* 40:1211-1215.

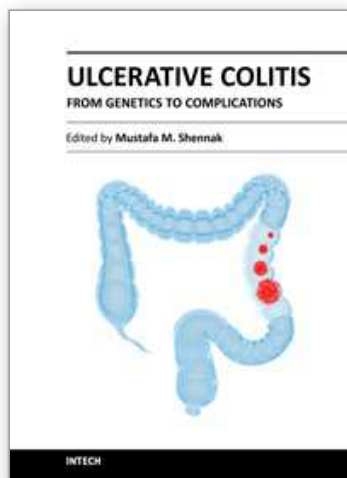
- Libioulle, C., E. Louis, S. Hansoul, C. Sandor, F. Farnir, D. Franchimont et al. 2007. Novel Crohn disease locus identified by genome-wide association maps to a gene desert on 5p13.1 and modulates expression of PTGER4. *PLoS Genet* 3:e58.
- McCarthy, M. I., G. R. Abecasis, L. R. Cardon, D. B. Goldstein, J. Little, J. P. Ioannidis, and J. N. Hirschhorn. 2008. Genome-wide association studies for complex traits: consensus, uncertainty and challenges. *Nat Rev Genet* 9:356-369.
- Myles, S., D. Davison, J. Barrett, M. Stoneking, and N. Timpson. 2008. Worldwide population differentiation at disease-associated SNPs. *BMC Med Genomics* 1:22.
- Nakagome, S., Y. Takeyama, S. Mano, S. Sakisaka, T. Matsui, S. Kawamura, and H. Oota. 2010. Population-specific susceptibility to Crohn's disease and ulcerative colitis; dominant and recessive relative risks in the Japanese population. *Ann Hum Genet* 74:126-136.
- Neel, J. V. 1962. Diabetes mellitus: a "thrifty" genotype rendered detrimental by "progress"? *Am J Hum Genet* 14:353-362.
- Ogura, Y., D. K. Bonen, N. Inohara, D. L. Nicolae, F. F. Chen, R. Ramos et al. 2001a. A frameshift mutation in NOD2 associated with susceptibility to Crohn's disease. *Nature* 411:603-606.
- Ogura, Y., N. Inohara, A. Benito, F. F. Chen, S. Yamaoka, and G. Nunez. 2001b. Nod2, a Nod1/Apaf-1 family member that is restricted to monocytes and activates NF-kappaB. *J Biol Chem* 276:4812-4818.
- Oota, H., A. J. Pakstis, B. Bonne-Tamir, D. Goldman, E. Grigorenko, S. L. Kajuna, N. J. Karoma, S. Kungulilo, R. B. Lu, K. Odunsi, F. Okonofua, O. V. Zhukova, J. R. Kidd, and K. K. Kidd. 2004. The evolution and population genetics of the ALDH2 locus: random genetic drift, selection, and low levels of recombination. *Ann Hum Genet* 68:93-109.
- Parkes, M., J. C. Barrett, N. J. Prescott, M. Tremelling, C. A. Anderson, S. A. Fisher et al. 2007. Sequence variants in the autophagy gene IRGM and multiple other replicating loci contribute to Crohn's disease susceptibility. *Nat Genet* 39:830-832.
- Pasvol, G., D. J. Weatherall, and R. J. Wilson. 1978. Cellular mechanism for the protective effect of haemoglobin S against *P. falciparum* malaria. *Nature* 274:701-703.
- Peltekova, V. D., R. F. Wintle, L. A. Rubin, C. I. Amos, Q. Huang, X. Gu, B. Newman, M. Van Oene, D. Cescon, G. Greenberg, A. M. Griffiths, P. H. St George-Hyslop, and K. A. Siminovitch. 2004. Functional variants of OCTN cation transporter genes are associated with Crohn disease. *Nat Genet* 36:471-475.
- Raelson, J. V., R. D. Little, A. Ruether, H. Fournier, B. Paquin, P. Van Eerdewegh et al. 2007. Genome-wide association study for Crohn's disease in the Quebec Founder Population identifies multiple validated disease loci. *Proc Natl Acad Sci U S A* 104:14747-14752.
- Rioux, J. D., M. J. Daly, M. S. Silverberg, K. Lindblad, H. Steinhart, Z. Cohen et al. 2001. Genetic variation in the 5q31 cytokine gene cluster confers susceptibility to Crohn disease. *Nat Genet* 29:223-228.
- Rioux, J. D., R. J. Xavier, K. D. Taylor, M. S. Silverberg, P. Goyette, A. Huett et al. 2007. Genome-wide association study identifies new susceptibility loci for Crohn disease and implicates autophagy in disease pathogenesis. *Nat Genet* 39:596-604.

- Risch, N., and K. Merikangas. 1996. The future of genetic studies of complex human diseases. *Science* 273:1516-1517.
- Risch, N. J. 2000. Searching for genetic determinants in the new millennium. *Nature* 405:847-856.
- Rosenberg, N. A., L. Huang, E. M. Jewett, Z. A. Szpiech, I. Jankovic, and M. Boehnke. 2010. Genome-wide association studies in diverse populations. *Nat Rev Genet* 11:356-366.
- Schaid, D. J., and S. S. Sommer. 1993. Genotype relative risks: methods for design and analysis of candidate-gene association studies. *Am J Hum Genet* 53:1114-1126.
- Shaw, M. H., N. Kamada, N. Warner, Y. G. Kim, and G. Nunez. 2011. The ever-expanding function of NOD2: autophagy, viral recognition, and T cell activation. *Trends Immunol* 32:73-79.
- The International HapMap Consortium. 2005. A haplotype map of the human genome. *Nature* 437:1299-1320.
- The International HapMap Consortium. 2007. A second generation human haplotype map of over 3.1 million SNPs. *Nature* 449:851-861.
- The Wellcom Trust Case Control Consortium. 2007. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* 447:661-678.
- Tishkoff, S. A., E. Dietzsch, W. Speed, A. J. Pakstis, J. R. Kidd, K. Cheung, B. Bonne-Tamir, A. S. Santachiara-Benerecetti, P. Moral, M. Krings, S. Paabo, N. Risch, T. Jenkins, and K. K. Kidd. 1996. Global patterns of linkage disequilibrium at the CD4 locus and modern human origins. *Science* 271:1380-1387.
- Tishkoff, S. A., and K. K. Kidd. 2004. Implications of biogeography of human populations for 'race' and medicine. *Nat Genet* 36:S21-27.
- Travassos, L. H., L. A. Carneiro, M. Ramjeet, S. Hussey, Y. G. Kim, J. G. Magalhaes et al. 2010. Nod1 and Nod2 direct autophagy by recruiting ATG16L1 to the plasma membrane at the site of bacterial entry. *Nat Immunol* 11:55-62.
- Xavier, R. J., and D. K. Podolsky. 2007. Unravelling the pathogenesis of inflammatory bowel disease. *Nature* 448:427-434.
- Yamaguchi-Kabata, Y., K. Nakazono, A. Takahashi, S. Saito, N. Hosono, M. Kubo, Y. Nakamura, and N. Kamatani. 2008. Japanese population structure, based on SNP genotypes from 7003 individuals compared to other ethnic groups: effects on population-based association studies. *Am J Hum Genet* 83:445-456.
- Yamazaki, K., D. McGovern, J. Ragoussis, M. Paolucci, H. Butler, D. Jewell et al. 2005. Single nucleotide polymorphisms in TNFSF15 confer susceptibility to Crohn's disease. *Hum Mol Genet* 14:3499-3506.
- Yamazaki, K., Y. Onouchi, M. Takazoe, M. Kubo, Y. Nakamura, and A. Hata. 2007. Association analysis of genetic variants in IL23R, ATG16L1 and 5p13.1 loci with Crohn's disease in Japanese patients. *J Hum Genet* 52:575-583.
- Yamazaki, K., A. Takahashi, M. Takazoe, M. Kubo, Y. Onouchi, A. Fujino, N. Kamatani, Y. Nakamura, and A. Hata. 2009. Positive association of genetic variants in the upstream region of NKX2-3 with Crohn's disease in Japanese patients. *Gut* 58:228-232.

- Yamazaki, K., M. Takazoe, T. Tanaka, T. Ichimori, S. Saito, A. Iida, Y. Onouchi, A. Hata, and Y. Nakamura. 2004. Association analysis of SLC22A4, SLC22A5 and DLG5 in Japanese patients with Crohn disease. *J Hum Genet* 49:664-668.
- Yamazaki, K., M. Takazoe, T. Tanaka, T. Kazumori, and Y. Nakamura. 2002. Absence of mutation in the NOD2/CARD15 gene among 483 Japanese patients with Crohn's disease. *J Hum Genet* 47:469-472.

IntechOpen

IntechOpen



Ulcerative Colitis from Genetics to Complications

Edited by Prof. Mustafa Shennak

ISBN 978-953-307-853-3

Hard cover, 222 pages

Publisher InTech

Published online 18, January, 2012

Published in print edition January, 2012

Ulcerative Colitis (UC) is a rapidly evolving medical field, and will continue to be very exiting in the next few decades. Although the underlying cause of this disease is still unknown, results in research dealing with various issues related to this disease are published every day. Chapters included in this book review the most recent literature on related advancements in regard to this chronic disease, which is controllable but not curable. Aspects like epidemiology, pathophysiology, genetics, incriminated etiologies, clinical aspects, complications, and disease management, including advancements in the diagnostic and therapeutic options, were documented by well known clinicians, researchers, and world wide authorities in their fields. This book on UC will be a valuable addition to each doctor's library interested in this subject, or for physicians dealing with patients suffering from this disease. Authors have also included figures and diagrams to depict their point, and to easily reach the minds of the readers in the simplest way.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Shigeki Nakagome and Hiroki Oota (2012). Evolutionary Insights into the “Population-Specificity” of the Genetic Factors Associated with Inflammatory Bowel Diseases, *Ulcerative Colitis from Genetics to Complications*, Prof. Mustafa Shennak (Ed.), ISBN: 978-953-307-853-3, InTech, Available from:
<http://www.intechopen.com/books/ulcerative-colitis-from-genetics-to-complications/evolutionary-insights-into-the-population-specificity-of-the-genetic-factors-associated-with-inflamm>

INTECH
open science | open minds

InTech Europe

University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China

Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821

© 2012 The Author(s). Licensee IntechOpen. This is an open access article distributed under the terms of the [Creative Commons Attribution 3.0 License](https://creativecommons.org/licenses/by/3.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

IntechOpen

IntechOpen