

Heritability in Inflammatory Bowel Disease: From the First Twin Study to Genome-Wide Association Studies

Hannah Gordon, BA, MBBS,* Frederik Trier Moller, MD,^{†,‡,§} Vibeke Andersen, MD, PhD,^{†,‡} and Marcus Harbord, BSc, MBBS, PhD*

Abstract: Since Tysk et al's pioneering analysis of the Swedish twin registry, twin and family studies continue to support a strong genetic basis of the inflammatory bowel diseases. The coefficient of heritability for siblings of inflammatory bowel disease probands is 25 to 42 for Crohn's disease and 4 to 15 for ulcerative colitis. Heritability estimates for Crohn's disease and ulcerative colitis from pooled twin studies are 0.75 and 0.67, respectively. However, this is at odds with the much lower heritability estimates from Genome-Wide Association Studies (GWAS). This "missing heritability" is likely due to shortfalls in both family studies and GWAS. The coefficient of heritability fails to account for familial shared environment. Heritability calculations from twin data are based on Falconer's method, with premises that are increasingly understood to be flawed. GWAS based heritability estimates may underestimate heritability due to incomplete linkage disequilibrium, and because some single nucleotide polymorphisms (SNPs) do not reach a level of significance to allow detection. SNPs missed by GWAS include common SNPs with low penetrance and rare SNPs with high penetrance. All methods of heritability estimation regard genetic and environmental variance as separate entities, although it is now understood that there is a complex multidirectional interplay between genetic and environmental factors mediated by the microbiota, the epigenome, and the innate and acquired immune systems. Due to the limitations of heritability estimates, it is unlikely that a true value for heritability will be reached. Further work aimed at quantifying the variance explained across GWAS, epigenome-wide, and microbiota-wide association studies will help to define factors leading to inflammatory bowel disease.

(*Inflamm Bowel Dis* 2015;21:1428–1434)

Key Words: heritability, missing heritability, twin studies, GWAS, inflammatory bowel disease

It has been over 25 years since Tysk et al's pioneering analysis of the Swedish twin registry. This important research included the first calculations of heritability from twin data, and inferred Crohn's disease (CD) to be largely due to genetic variance, with ulcerative colitis (UC) underpinned by a lesser yet significant genetic basis.

Subsequent twin and family studies have continued to support a strong genetic component to IBD incidence, although perhaps to a lesser degree than seen amongst Tysk's cohort. Having a first-degree relative with CD or UC confers a greater risk than any known environmental factor.² However, the rate at which incidence has risen worldwide in the past century,³ and more recently in Asia,⁴ significantly exceeds that which can be explained by genetic drift.

Heritability is defined as the proportion of phenotypic variance that can be attributed to genetic variance.⁵ Various methods have been used to calculate heritability in IBD. Twin studies frequently use the classical twin design. Estimations of the coefficient of heritability from family studies are used to infer the increased risk in family members. More recently, estimation of the effects of single nucleotide polymorphisms (SNPs) identified from Genome-Wide Association Studies (GWAS) are used to calculate heritability.

These methods yield strikingly different results. In particular, results of GWAS account for less than 50% of heritability estimated by twin studies.⁶ This phenomenon is known as missing or hidden heritability.

This article reviews what family studies, twin studies, and GWAS have taught us about heritability since Tysk's initial analysis. Furthermore, the assumptions and challenges of each method are evaluated, shedding light on why heritability estimates are so variable. Finally, we propose a model for the complex genetic–environmental interactions which mediate IBD onset, and discuss how more accurate evaluation of environmental and genetic variance could be achieved in the future.

FAMILY STUDIES; LESSONS AND LIMITATIONS

Familial clustering of CD and UC was noticed shortly after both diseases were described in the 1930s.^{7,8} A family history of IBD is reported by 5% to 16% of patients with CD and 8% to 14% patients with UC.^{2,9,10} Family studies continue to show clustering

Received for publication January 2, 2015; Accepted February 9, 2015.

From the *Department of Gastroenterology, Chelsea and Westminster Hospital, London, United Kingdom; [†]Organ Center, Hospital of Southern Jutland, Aabenraa, Denmark; [‡]Institute of Regional Health Research-Center Sønderjylland, University of Southern Denmark, Odense, Denmark; and [§]Department of Epidemiology Research, Statens Serum Institut, Copenhagen, Denmark.

Supported in part by an unrestricted grant from the Jesse and Thomas Tam charitable foundation.

The authors have no conflicts of interest to disclose.

Reprints: Hannah Gordon, BA, MBBS, Department of Gastroenterology, Chelsea and Westminster Hospital, 369 Fulham Road, London SW10 9NH, United Kingdom (e-mail: hannahgordon@doctors.org.uk).

Copyright © 2015 Crohn's & Colitis Foundation of America, Inc.

DOI 10.1097/MIB.0000000000000393

Published online 17 April 2015.

of disease, most prominent in close relatives. However, increased incidence has also been found in second- and third-degree relatives.¹¹ Having 1 or more affected first degree relatives is still the greatest identified risk factor,¹² with the lifetime risk to offspring of 2 parents with IBD exceeding 30%.¹³

Several studies have shown correlations between familial disease incidence, disease location, severity,¹⁴ and extraluminal manifestations. Patients with IBD with a positive family history also tend to present at a younger age, with CD sufferers more likely to have small bowel disease.¹⁵

Familial clustering varies with geographic location and ethnicity. American studies show greater familial concordance than Scandinavian cohorts. Ashkenazi Jews show the greatest rates of family concordance,¹⁶ and Hispanic and Asian populations the lowest.¹⁷ There are also low rates of disease among first-degree relatives of Asian IBD sufferers within the United Kingdom.² The genetic basis of disease also varies; CARD15 is associated with CD in the West and is particularly well represented among Ashkenazi Jews.¹⁸ However, CARD15 is not associated with IBD in Asian populations; instead, TNF-SF15 polymorphisms are associated with CD and TNF-308 and CTLA-4 polymorphisms associated with UC.¹⁹ Although this genetic variance may reflect differing genetic susceptibility, it cannot explain the rapid rise of IBD in the East, and indeed worldwide during the past century.

Perhaps, one of the strengths of family studies has been their ability to discover rare genes with strong penetrance. For example, family studies of children who develop IBD within the first 2 years of life have led to the identification of homozygous mutations in the interleukin-10 receptor and interleukin-10 genes in approximately 1/3 of all very early-onset IBD cases.^{20–22} This phenotype is associated with a very strong family history. It is also more common in consanguineous families, in whom case reports suggest Mendelian recessive inheritance.^{23,24}

The coefficient of heritability (λ) is a measure of the odds ratio of IBD prevalence in a defined family member of the proband and the population prevalence. A high coefficient of heritability is often taken to infer a greater genetic influence on phenotypic variance. It is most often measured for siblings (λ_s). λ_s is 25 to 42 for CD and 4 to 15 for UC,²⁵ depending on the study population. This is higher than the coefficient of heritability in many other complex disorders, including schizophrenia and type 1 diabetes.²⁶ However, a major shortfall of the coefficient of heritability is that it does not account for the more similar environment shared by family members. A key strength of Tysk's twin analysis, and subsequent twin studies which have followed, is that comparison of monozygotic and dizygotic twins allows a degree of control for environmental factors.

HERITABILITY AND TWIN STUDIES

Since Tysk's original analysis of the Swedish twin literature, several twin studies have been undertaken. One is a review of data from the Swedish registry, and other major IBD twin studies from Danish, British, and German populations. Sample

sizes range from 102 to 250 twin pairs.^{27–31} The German and U.K. subjects were recruited from patient support groups, whereas Scandinavian cohorts were recruited from national twin registries.

There are various twin study designs, which have been used to calculate or infer heritability of disease. The classical twin study compares concordance rates between monozygotic (MZ) and dizygotic (DZ) twin pairs. Other methodology includes adoption studies, longitudinal studies, and studies of the offspring of monozygotic twins.³² The classical design has been most widely used to infer or calculate heritability in IBD.

The data from twin studies are most often presented as pair and proband concordance in MZ and DZ groups. Proband concordance is considered the more robust measure and takes into account how both concordant twins were identified.¹ This method is of particular value when a single twin within an existing twin registry is used to search for cases (Fig. 1).³³

The range of concordance rates in published IBD twin studies is illustrated in Table 1.^{29–32,34} It can be seen that MZ concordance is consistently higher than DZ concordance, inferring heritability in both diseases. This is particularly true of CD. However, lifelong concordance does not approach 100%, emphasizing the importance of environmental factors. A Norwegian population-based study reported DZ twin concordance to be the double that of other siblings,³⁵ possibly pointing to an importance of early environmental factors more likely to be shared between twins, including intrauterine environment.

Concordance rates vary significantly across the European twin cohorts, which could be due to study design heterogeneity. Previous twin research has suggested a disproportionately greater response to questionnaires from MZ twins.³⁶ Furthermore, the inception ages and follow-up duration vary between populations; young discordant twins may become concordant later in life. Recently, methods have been developed that calculate heritability from twin data while accounting for censoring and competing risk.³⁷ However, such methods have not yet been used in IBD.

Although the majority of twin literature focuses on concordance, calculations of heritability have been made from twin studies. Models are most often based on Falconer's method,³⁸ summarized in Figure 2. A key assumption is that environmental factors are equally similar between MZ and DZ twin pairs. As such differences in concordance between the groups and knowledge of the underlying population, prevalence can be used to calculate genetic variance. High MZ concordance is itself not necessarily proof of high heritability (e.g., MZ twins will be highly concordant for history of chickenpox); it is the difference between concordance of MZ and DZ pairs which is paramount.

Falconer's method requires calculation of the intraclass correlation for each group. This is the degree to which individuals within a group resemble each other in terms of a trait. In this case, the groups are MZ and DZ twin pairs. Traits can be directly correlated when comparing traits which are measured on a linear scale and known to follow a normal distribution.³³ Examples of such traits include height or weight. However, when the trait is binary, such as the presence or absence of IBD, it is assumed that

Pair concordance represents the percentage of all twin pairs within the sample who are concordant for disease.

$$\text{Pair concordance} = \frac{C}{C+D}$$

C = number of concordant twin pairs

D = number of discordant twin pairs

Proband concordance can be calculated when a known cohort of twins is assessed for cases with disease. This method takes note of how concordant pairs are discovered; namely whether both cases are found separately, or whether the second diagnosis is made following communication with an index case.

Proband Concordance

$$= \frac{2(C2) + C1}{2(C2) + C1 + D}$$

C1 = Concordant pairs discovered following interview with index case

C2 = Concordant twins both discovered as index cases

D = Discordant twin pairs

FIGURE 1. Pair and proband concordance.

there is an underlying liability to disease which itself is normally distributed. Once a certain threshold of liability is reached, the trait becomes present. Reich's method can be used to calculate intraclass correlations from dichotomous data; details of this can be found within Reich's workings.³⁹

Chen et al applied Falconer's equation with Reich's method to calculate heritability from the pooled results of the IBD twin studies.⁴⁰ The heritability of CD and UC were found to be 0.75

TABLE 1. Range of Concordance from Published IBD Twin Studies

	CD, %	UC, %
Monozygotic	20–55	6.3–17
Dizygotic	0–3.6	0–6.3

and 0.67, respectively. Pooling data provided sufficient sample size for satisfactory analysis. However, the validity of pooled data is limited by differences in twin study methodology.

The assumptions underpinning Falconer's equation (Fig. 2) are likely to be an over simplification of the true genetic and environmental interplay. As discussed, the model assumes an underlying scale of liability to disease, which itself is determined by normally distributed genetic and environmental factors. This assumption cannot be accurately tested. Mathematical modeling suggests that even a small deviation from this assumption significantly influences results.⁴¹ Thus, if the underlying variation does not follow a normal distribution, the true heritability of liability may be significantly distorted.

Falconer's model calculates narrow sense as opposed to broad sense heritability. This assumes exclusively additive genetic effects, such that all genes coding for the CD or UC phenotype "add up" and do not interact with each other.³³ This does not occur in complex biological systems and does not account for epistasis or genetic dominance. Epistasis describes the interaction of genes from separate loci⁴² and dominance the relationship between alleles of a single gene.

Estimations of heritability from twin studies assume that MZ twins share 100% segregating genes, whereas DZ twins share 50%. Neither assumption is accurate. The relatedness of DZ twins is determined by the recombination of genes during prophase 1 of meiosis; the proportion of alike genetic material each sibling receives varies. Genome-wide microsatellite marker data suggest that the true value lies between 42% and 60%.⁴³ The assumption that MZ twins are genetically identical is also an over simplification. Copy number variations (genomic duplications and deletions⁴⁴) exist between monozygotic twin pairs.⁴⁵ This has not been studied extensively in twin pairs with IBD but has been implicated in twins discordant for schizophrenia.⁴⁶

The classical twin study also assumes MZ and DZ twins share a similar environment. This assumption, known as the equal environment assumption, is flawed. Even the intrauterine environment differs between MZ and DZ pairs. For example, DZ twins always have a separate placenta, whereas 30% of MZ twins are dichorionic.⁴⁷ Second, DZ twins may be of a different gender, thus sex bias will only influence environmental differences within the DZ group. Finally, adoption studies have shown that complex behavioral traits (e.g., smoking) themselves have a degree of heritability.⁴⁸ Thus, MZ twins may self-select a more similar environment. The extent to which small deviations from an

Falconer's Equation

$$h^2 = 2(r_{MZ} - r_{DZ})$$

h^2 = narrow sense heritability

r_{MZ} = intra-class correlation of monozygotic twins

r_{DZ} = intra-class correlation of dizygotic twins

Assumptions when estimating heritability from twin studies:

1. The presence or absence of disease is determined by underlying genetic and environmental factors which themselves are normally distributed (heritability of liability)
2. All genetic influences are additive (only narrow sense heritability is estimated)
3. Monozygotic twins share 100% segregating genes
4. Dizygotic twins share on average 50% segregating genes
5. Monozygotic and dizygotic twin pairs have an equally similar environment
6. Genetic and Environmental variations act independently of one another

FIGURE 2. Falconer's equation and the assumptions underpinning this model.

equal environment influence heritability is debatable. Although geneticists often refute the equal environment assumptions for the reasons given above,⁴⁹ most mathematical models suggest the bias it confers is modest.⁵⁰

GWAS AND MISSING HERITABILITY

In recent years, GWAS have been used to calculate heritability, yielding strikingly different results from twin literature. SNPs are variations in DNA sequence, which occur significantly more frequently in populations with a trait than populations without the trait. In GWAS studies, SNPs associated with the outcome of interest are identified. Heritability can be calculated from GWAS studies from the additive genetic variance attributable to all identified SNPs—this is known as the SNP effects. Subsequently, regression models are used to estimate narrow sense heritability. An overview of the mathematical models and a run through of these methods using data from GWAS in CD are provided in Hong Lee et al's 2011 article.⁵¹

A meta-analysis of GWAS and immuno-chip data undertaken by the International IBD Genetics Consortium has identified 163 SNPs associated with IBD.⁵² One hundred ten SNPs have associations with both disease phenotypes. This is corroborative with pedigree studies, which have long suggested overlapping inheritance. The estimates of heritability of liability for CD and UC, calculated from imputed SNP effects, are 0.37 and 0.27, respectively; this is less than half of the 0.75 and 0.67 calculated from pooled twin data.⁴⁰

This phenomenon, known as missing or hidden heritability, is not exclusive to IBD. Schizophrenia and autism both have heritability estimates from twin studies in the range of 0.8⁵³ and 0.9,⁵⁴ respectively; however, SNPs identified by GWAS account

for only a small fraction. Although the twin literature may overestimate heritability, biasing the “denominator” of the missing heritability calculation, there are also limitations of the methods used when calculating heritability from GWAS. These are reviewed below.

Linkage disequilibrium refers to the nonrandom association of alleles at 2 or more linked genetic loci. LD occurs when loci are passed on together from parents to offspring. If significant genes lie within regions of the genome not directly probed, it is thought that the GWAS will identify SNPs in LD with these genes. However, if LD is incomplete, SNPs may not always be passed on with the coding genes, and as such coding genes will not always be detected by GWAS.⁵⁵

GWAS do not cover every polymorphism of the genome, and approximately 30% of genes are not reviewed.⁵⁶ In addition to this, using GWAS to enumerate, the total genetic burden causing disease will miss SNPs that do not reach the somewhat arbitrary level of statistical significance. As such, both rare variants with high penetrance and common variants with low penetrance may not be detected.⁵⁷

Studies searching for rare variants among coding exons of 25 GWAS identified risk genes for autoimmune disease suggest that the role of rare variants is minimal.⁵⁷⁻⁵⁹ It is more likely that GWAS miss a multitude of common alleles with a penetrance that is too low to be detected.^{60,61} The practical value of discovering SNPs with very small effects is debatable. However, in summation, they may account for some of the missing heritability. Unfortunately, without their identification, it is impossible to calculate their contribution.

The GWAS heritability estimate may also be falsely low when causal variants are poorly tagged by any single GWAS marker or if multiple independent causal variants exist at the identified GWAS locus.⁶² Gusev et al proposed to make use of all observed markers in a variance-components analysis to estimate the

total contribution of all typed markers in the sample. Their methods did not require individual markers to be genome-wide significant. When applied to known GWAS loci, this method yielded significant increases in the liability-scale heritability for CD and UC.⁶²

As with the twin models, the assumption of exclusively additive genetic variance may well be incorrect.⁶³ Simulated data have suggested epistasis has less impact on genetic variance in populations of unrelated individuals; as such if epistasis is present, twin studies will be more biased than estimations of SNP effects. This may contribute to the heritability gap.

The larger GWAS studies involve meta-analysis resulting in very large cohorts. Heterogeneity within the cohorts in terms of both SNPs and environmental triggers may prevent the identification of SNPs. Some SNPs will only be detected in settings with the relevant exposures, thus different exposures across the cohorts will lead to different genes detected. Thus, when analyzing data from different populations, different environmental factors may bias estimations of heritability. Another limitation is that GWAS will not properly detect structural changes in genetic architecture such as copy number variations, deletions, insertions, or epigenetic variations unless they are in complete LD with 1 or more GWAS marker.

GENE–ENVIRONMENT INTERACTIONS AND INFLUENCE ON HERITABILITY

The underlying assumption that genetic and environmental factors act independently of one another still underpins the mathematical models used to estimate heritability. However, there is increasing evidence that this underlying premise is false, or at least an oversimplification.

The true interaction between genes and the environment most likely involves a complex multidirectional interplay, mediated by the epigenome, microbiota, and the innate and acquired immune system. Environmental factors themselves may directly influence risks attributable to specific SNPs; one study showed SNPs from the IBD4 locus only reach genome-wide significance in smokers.⁶⁴ Smoking is known to influence the microbiota as are diet and antibiotic use. Epigenetic change may alter the function of the innate immune system. These are but a few of the possible interactions between genes and environment, which are summarized in Figure 3.^{65–71}

CONCLUSIONS AND FUTURE DIRECTION

Since Tysk's first analysis of the twin literature, further twin studies and family studies continue to support a genetic component to the IBD phenotype. GWAS studies further support this, with identification of 163 SNPs associated with CD and UC. However, quantitative estimations of heritability are far from straight forward, and a more complete understanding of gene–environmental interactions makes this more challenging. The way in which results from studies are presented varies considerably; heritability coefficients from family studies can neither be compared with concordance from twin studies nor variance calculated from GWAS without further statistical

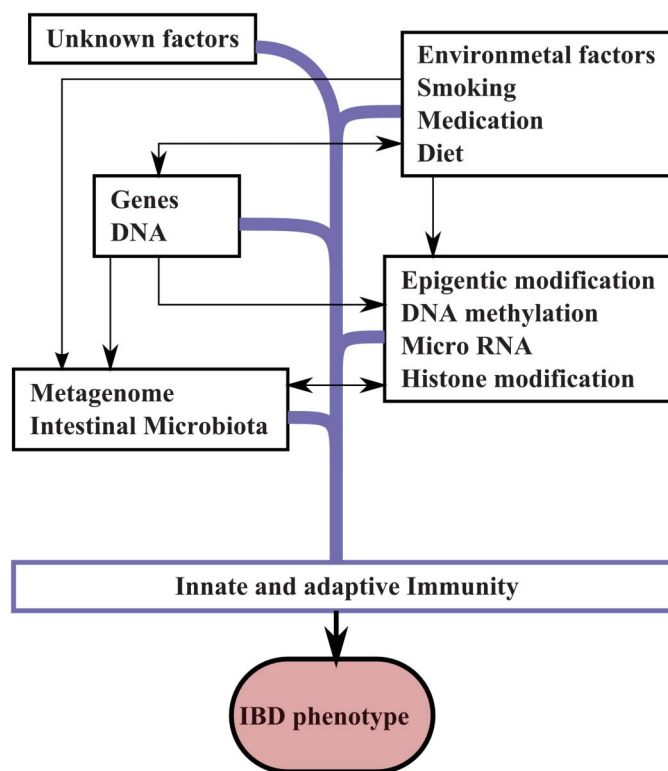


FIGURE 3. The complexity of gene–exposome interactions associated with the IBD phenotype.

manipulation. All of the models include unverified assumptions. Thus, it is likely that a true quantitative value of heritability in IBD may never be reached.

Although an absolute measurement of total heritability may not be possible, this does not detract from the achievements of family and twin studies over the years. They have helped to identify genetic variants associated with IBD phenotype, and may continue to identify SNPs of low frequency but high penetrance. Indeed, studies of multiplex families followed by LD mapping led to the discovery of the first IBD-related SNPs⁷² including NOD2.⁷³ The inferred high heritability from twin and family literature prompted this work, leading to the choice of CD and UC for GWAS interrogation, which has subsequently uncovered 163 SNPs for future pathway analysis.

Coupled with new omics techniques, information from family studies and GWAS continue to enhance our knowledge. The traditional pathway of genes and environment leading to IBD phenotype has been challenged—the role of the microbiota linking the environment to phenotype may prove particularly important as it can be manipulated. At present, large scale studies of the genome, epigenome, microbiota, and metabolome within the same individuals with IBD are underway.⁷⁴ This may facilitate understanding the relative importance of each factor. Studies of inception cohorts have identified changes to the microbiota, which are present immediately after diagnosis.⁶⁹ The GEM project is using our knowledge of familial clustering in an attempt to

capture a “before” and “after” snap shot of the epigenome and microbiota of healthy individuals who are at high risk of developing CD.⁷⁵ Discordant twins continue to provide a natural avenue for paired analysis; and may prove particularly important for future epigenetics work. Collaboration between twin registries will facilitate larger sample sizes for this work to progress.

Finally, geneticists are crucial to understanding the relative weighting of the factors involved in pathogenesis. It is likely that the mathematical models of variance and heritability will continue to develop such that the confounding factors described in this article are better assessed. In turn, this will shed further light upon their relative importance and guide the direction of future research efforts. Although heritability may not be the answer, quantifying the variance explained across GWAS, epigenome-wide, and microbiota-wide association studies will help to bring us closer to understanding those factors that lead to IBD.

REFERENCES

- Tysk C, Lindberg E, Järnerot G, et al. 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;29:990–996.
- Probert CSJ, Jayanthi V, Hughes AO, et al. Prevalence and family risk of ulcerative colitis and Crohn's disease: an epidemiological study among Europeans and South Asians in Leicestershire. *Gut*. 1993;34:1547–1551.
- Molodecky NA, Soon IS, Rabi DM, et al. Increasing incidence and prevalence of the inflammatory bowel diseases with time, based on systematic review. *Gastroenterology*. 2012;142:46–54.e42; quiz e30.
- Kamm MA, De Cruz PP, Chan FK, et al. Inflammatory bowel disease in Asia: a systematic review. *J Gastroenterol Hepatol*. 2012;27:1266–1280.
- Wray N, Visscher P. Estimating trait heritability. *Nat Educ*. 2008;1:29.
- Franke A, McGovern DPB, Barrett JC, et al. Meta-analysis increases to 71 the tally of confirmed Crohn's disease susceptibility loci. *Nat Genet*. 2010;42:1118–1125.
- Spriggs EI. Chronic ulceration of the colon. *Q J Med*. 1934;27:549.
- Kirsner JB, Spencer JA. Family occurrences of ulcerative colitis, regional enteritis, and ileocolitis. *Ann Intern Med*. 1963;59:133–144.
- Halme L, Turunen U, Heliö T, et al. Familial and sporadic inflammatory bowel disease: comparison of clinical features and serological markers in a genetically homogeneous population. *Scand J Gastroenterol*. 2002;37:692–698.
- Yang H, McElree C, Roth MP, et al. Familial empirical risks for inflammatory bowel disease: differences between Jews and non-Jews. *Gut*. 1993;34:517–524.
- Trier Moller F, Anderson V, Jess T. DOP021 Familial risk of inflammatory bowel disease: a population-based cohort study 1977–2011. *JCC*. 2014;8(Suppl 1):S24–S25.
- Peeters M, Nevens H, Baert F, et al. Familial aggregation in Crohn's disease: increased age-adjusted risk and concordance in clinical characteristics. *Gastroenterology*. 1996;111:597–603.
- Halme L, Paavola-Sakki P, Turunen U, et al. Family and twin studies in inflammatory bowel disease. *World J Gastroenterol*. 2006;12:3668–3672.
- Satsangi J, Grootsholten C, Holt H, et al. Clinical patterns of familial inflammatory bowel disease. *Gut*. 1996;38:738–741.
- Carbannel F, Macaigne G, Beaugerie L, et al. Crohn's disease severity in familial and sporadic cases. *Gut*. 1999;44:91–95.
- Rotter JI, Yang H, Shohat T. Genetic complexities of inflammatory bowel disease and its distribution among the Jewish people. In: Bonne-Tamir B, Adam A, eds. *Genetic Diversity Among Jews: Disease and Markers at the DNA Level*. Oxford, United Kingdom: Oxford University Press; 1992:395–411.
- Kurata JH, Kantor-Fish S, Frankl H, et al. Crohn's disease among ethnic groups in a large health maintenance organization. *Gastroenterology*. 1992;102:1940–1948.
- Bonen DK, Ogura Y, Nicolae DL, et al. Crohn's disease-associated *NOD2* variants share a signalling defect in response to lipopolysaccharide and peptidoglycan. *Gastroenterology*. 2003;124:140–146.
- Ng SC, Tsoi KK, Kamm MA, et al. Genetics of inflammatory bowel disease in Asia: systematic review and meta-analysis. *Inflamm Bowel Dis*. 2012;18:1164–1176.
- Glocker EO, Kotlarz D, Boztug K, et al. Inflammatory bowel disease and mutations affecting the interleukin-10 receptor. *N Engl J Med*. 2009;361:2033–2045.
- Glocker EO, Frede N, Perro M, et al. Infant colitis—it's in the genes. *Lancet*. 2010;376:1272.
- Lee CH, Hsu P, Nanan B, et al. Novel de novo mutations of the interleukin-10 receptor gene lead to infantile onset inflammatory bowel disease. *J Crohns Colitis*. 2014;8:1551–1556.
- Megarbane A, Sayad R. Early lethal autosomal recessive enterocolitis: report of a second family. *Clin Genet*. 2007;71:89–90.
- Fried K, Vure E. A lethal autosomal recessive enterocolitis of early infancy. *Clin Genet*. 1974;6:195–196.
- Russell RK, Satsangi J. IBD: a family affair. *Best Pract Res Clin Gastroenterol*. 2004;18:525–539.
- Orholm M, Munkholm P, Langholz E, et al. Familial occurrence of inflammatory bowel disease. *N Engl J Med*. 1991;324:84–88.
- Orholm M, Binder V, Sorensen TI, et al. Concordance of inflammatory bowel disease among Danish twins. Results of a nationwide study. *Scand J Gastroenterol*. 2000;35:1075–1081.
- Jess T, Riis L, Jespersgaard C, et al. Disease concordance, zygosity, and *NOD2/CARD15* status: follow-up of a population-based cohort of Danish twins with inflammatory bowel disease. *Am J Gastroenterol*. 2005;100:2486–2492.
- Spehlmann ME, Begun AZ, Burghardt J, et al. Epidemiology of inflammatory bowel disease in a German twin cohort: results of a nationwide study. *Inflamm Bowel Dis*. 2008;14:968–976.
- 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;124:1767–1773.
- Ng SC, Woodrow S, Patel N, et al. Role of genetic and environmental factors in British twins with inflammatory bowel disease. *Inflamm Bowel Dis*. 2012;18:725–736.
- van Dongen J, Slagboom PE, Draisma HH, et al. The continuing value of twin studies in the omics era. *Nat Rev Genet*. 2012;13:640–653.
- Verweij KJ, Mosing MA, Zietsch BP, et al. Estimating heritability from twin studies. *Methods Mol Biol*. 2012;850:151–170.
- Brand SR. Update on the heritability of Inflammatory Bowel Disease: the importance of twin studies. *Inflamm Bowel Dis*. 2011;17:1–5.
- Bengtson MB, Aamodt G, Vatn MH, et al. Concordance for IBD among twins compared to ordinary siblings—a Norwegian population based study. *J Crohns Colitis*. 2010;4:312–318.
- Jablon S, Neel JV, Gershowitz H, et al. The NAS-NRC twin panel: methods of construction of the panel, zygosity diagnosis, and proposed use. *Am J Hum Genet*. 1967;19:133–161.
- Scheike TH, Holst KK, Hjelmberg JB. Estimating twin concordance for bivariate competing risks twin data. *Stat Med*. 2014;33:1193–1204.
- Falconer DS. The inheritance of liability to certain diseases, estimated from the incidence among relatives. *Ann Hum Genet*. 1966;29:51–76.
- Reich T, James JW, Morris CA. The use of multiple thresholds in determining the mode of transmission of semi-continuous traits. *Ann Hum Genet*. 1972;36:163–184.
- Chen GB, Lee SH, Brion MJ, et al; International IBD Genetics Consortium. Estimation and partitioning of (co)heritability of inflammatory bowel disease from GWAS and immunochip data. *Hum Mol Genet*. 2014;23:4710–4720.
- Benckek PH, Morris NJ. How meaningful are heritability estimates of liability? *Hum Genet*. 2013;132:1351–1360.
- Cordell HJ. Epistasis: what it means, what it doesn't mean, and statistical methods to detect it in humans. *Hum Mol Genet*. 2002;11:2463–2468.
- Visscher PM, Macgregor S, Benyamina B, et al. Genome partitioning of genetic variation for height from 11,214 sibling pairs. *Am J Med Genet*. 2007;81:1104–1110.
- Zhang F, Gu W, Hurler ME, et al. Copy number variation in human health, disease, and evolution. *Annu Rev Genomics Hum Genet*. 2009;10:451–481.

45. Ehli EA, Abdellaoui A, Hu Y, et al. De novo and inherited CNVs in MZ twin pairs selected for discordance and concordance on attention problems. *Eur J Hum Genet.* 2012;11:1037–1043.
46. Castellani CA, Awamleh Z, Melka MG, et al. Copy number variation distribution in six monozygotic twin pairs discordant for schizophrenia. *Twin Res Hum Genet.* 2014;17:108–120.
47. Curran, Mark (2005-11-02). "Twinning." Focus Information Technology. Retrieved 2008-10-10.
48. McGue M, Bouchard TJ Jr. Genetic and environmental influences on human behavioural differences. *Annu Rev Neurosci.* 1998;21:1–24.
49. Richardson K, Norgate S. The equal environments assumption of classical twin studies may not hold. *Br J Educ Psychol.* 2005;75:339–350.
50. Derks EM, Dolan CV, Boomsma DI. A test of the equal environment assumption (EEA) in multivariate twin studies. *Twin Res Hum Genet.* 2006;9:403–411.
51. Lee SH, Wray NR, Goddard ME, et al. Estimating missing heritability for disease from genome-wide association studies. *Am J Hum Genet.* 2011; 88:294–305.
52. Jostins L, Ripke S, Weersma RK, et al. Host-microbe interactions have shaped the genetic architecture of inflammatory bowel disease. *Nature.* 2012;491:119–124.
53. Sullivan PF. The genetics of schizophrenia. *PLoS Med.* 2005;2:e212.
54. Freitag CM. The genetics of autistic disorders and its clinical relevance: a review of the literature. *Mol Psychiatry.* 2007;12:2–22.
55. Goddard ME, Wray NR, Verbyla K, et al. Estimating effects and making predictions from genome-wide marker data. *Stat Sci.* 2009;24:517–529.
56. Lee JC, Parkes M. Genome-wide association studies and Crohn's disease. *Brief Funct Genomics.* 2011;10:71–76.
57. Bodmer W, Bonilla C. Common and rare variants in multifactorial susceptibility to common diseases. *Nat Genet.* 2008;40:695–701.
58. Hunt KA, Mistry V, Bockett NA, et al. Negligible impact of rare autoimmune-locus coding-region variants on missing heritability. *Nature.* 2013;498:232–235.
59. Wray N, Purcell S, Visscher P. Synthetic associations created by rare variants do not explain most GWAS results. *PLoS Biol.* 2011;9:e1000579.
60. Park JH, Wacholder S, Gail MH, et al. Estimation of effect size distribution from genome-wide association studies and implications for future discoveries. *Nat Genet.* 2010;42:570–575.
61. Yang J, Benyamin B, McEvoy BP, et al. Common SNPs explain a large proportion of the heritability for human height. *Nat Genet.* 2010;42:565–569.
62. Gusev A, Bhatia G, Zaitlen N, et al. Quantifying missing heritability at known GWAS loci. *PLoS Genet.* 2013;9:e1003993.
63. Zuk O, Hechter E, Sunyaev SR, et al. The mystery of missing heritability: genetic interactions create phantom heritability. *Proc Natl Acad Sci U S A.* 2012;109:1193–1198.
64. Pierik M, Yang H, Barmada MM, et al. The IBD international genetics consortium provides further evidence for linkage to IBD4 and shows gene-environment interaction. *Inflamm Bowel Dis.* 2005;11:1–7.
65. Molodecky NA, Kaplan GG. Environmental risk factors for inflammatory bowel disease. *Gastroenterol Hepatol (N Y).* 2010;6:339–346.
66. Nimmo ER, Prendergast JG, Aldhous MC, et al. Genome-wide methylation profiling in Crohn's disease identifies altered epigenetic regulation of key host defense mechanisms including the Th17 pathway. *Inflamm Bowel Dis.* 2012;18:889–899.
67. Wu F, Guo NJ, Tian H, et al. Peripheral blood microRNAs distinguish active ulcerative colitis and Crohn's disease. *Inflamm Bowel Dis.* 2011;17: 241–250.
68. Tsaprouni L, Ito K, Powell J, et al. Differential patterns of histone acetylation in inflammatory bowel diseases. *J Inflamm (Lond).* 2011;8:1.
69. Gevers D, Kugathasan S, Denson Lee A, et al. The treatment-naive microbiome in new-onset Crohn's disease. *Cell Host Microbe.* 2014;15:382–392.
70. Hooper LV, Littman DR, Macpherson AJ. Interactions between the microbiota and the immune system. *Science.* 2012;336:1268–1273.
71. Baumgart DC, Carding SR. Inflammatory bowel disease: cause and immunobiology. *Lancet.* 2007;369:1627–1640.
72. Satsangi J, Jewell DP, Bell JI. The genetics of inflammatory bowel disease. *Gut.* 1997;40:572–574.
73. 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.
74. An European consortium to boost our understanding of inflammatory bowel disease. Available at: www.ibdcharacter.eu/. Accessed December 12, 2014.
75. Crohn's and colitis Canada inflammatory bowel disease GEM project. Available at: www.gemproject.ca. Accessed December 20, 2014.