Dissecting the genetics of Inflammatory Bowel Disease

Heather Elding

January 2015

Thesis submitted for the degree of

Doctor of Philosophy

to

Research Department of Genetics, Evolution and Environment

University College London (UCL)

I, Heather Elding confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Abstract

Inflammatory Bowel Disease (IBD) can be classified into two main subtypes: Crohn's Disease (CD) and Ulcerative Colitis (UC). The aim of this study is to identify the genetic contribution to the susceptibility to IBD. In the first part of the study, I focused on Crohn's Disease, the subtype that shows the greatest heritability. Using both pooled and sub-phenotype data, followed by replication, the results reveal substantial genetic heterogeneity and the total number of confirmed CD susceptibility loci was increased from 71 published by others to 200. This was achieved by analyzing the data using a multi-marker approach and high-resolution genetic maps in Linkage Disequilibrium Units. In the second part of the study, I focused on Ulcerative Colitis. The results show that although UC has a lower reported heritability, many loci were also found for Ulcerative Colitis. Some of these overlap with those found for Crohn's Disease.

Acknowledgements

I would firstly like to thank my primary Ph.D supervisor Dr. Nikolas Maniatis for providing me with this opportunity and for his expert guidance and significant support throughout my time at UCL. Dr. Nikolas Maniatis's challenging ideas, creativity, and moral support have greatly helped me in shaping my thoughts as a researcher and for successfully carrying out my PhD research. I would also deeply like to thank my secondary Ph.D supervisor Professor Dallas M. Swallow for her invaluable and exemplary scientific expertise and constant support throughout my Ph.D. Professor Dallas M. Swallow has been, and will be, a great inspiration on both a professional and personal level and I have learnt a lot and made significant progress under her supervision, both as a scientist and as an individual. I would also like to thank Dr. Winston Lau for his patience at always taking time to teach me the novel skills and tools that were crucial for carrying out my Ph.D and for always being there to answer my questions.

I would like to thank Dr. Adrien Rieux for providing valuable feedback on my Ph.D Thesis, Dr. Toby Andrew, Dr. Doug Speed and Dr. Kaustubh Adhikari for their opinions on certains aspects of the analysis carried out in this thesis work and Karim Boustani who, as a summer student, has contributed with invaluable work.

I would also like to thank my parents Josette and Mario Elding and my sister Davinia Elding for supporting me throughout this journey. I would like to thank my aunt Maria Jordanov for always making sure I always had a plain-sailing stay in London and for her support and my aunt Anna Christie for always being there in time of need.

I would sincerely like to show my great appreciation to the Departmental Graduate Tutors Dr. Julia Day, Professor Kevin Fowler and Dr. Lazaros Foukas for their great help and understanding as well as their emotional support during very difficult times and life obstacles that I have encountered during my Ph.D.

Last, but definitely not least, I would like to thank all my friends, particularly Arjunan Rajasingam, Leanne Grech, Adrien Rieux, Florent Lassalle, Pascale Gerbault, Anke Liebert, my colleagues and friends at UCL Genetics Institute and all the people in my office in the Darwin Building for making my time at UCL an unforgettable happy time and for always being there for me.

Table of Contents

Table of Figures

CHAPTER I. GENERAL INTRODUCTION

I. INFLAMMATORY BOWEL DISEASE: THE STORY OF TWO DISEASES.

Inflammatory Bowel Disease (IBD) is an immunologically mediated relapsing inflammatory condition of unknown causes. IBD can be subdivided into two sub-types: Crohn's Disease (CD) and Ulcerative Colitis (UC), as illustrated in Table 1. In the second half of the twentieth century, the incidence of IBD has been on the increase in the developed western world (Sartor, 2006). This suggests the presence of a strong environmental contribution to the aetiology of IBD. Therefore, why study the genetics of IBD? Evidence for a genetic contribution comes from increased concordance in Monozygotic twins (MZ) (~50% for CD and ~10% for UC) in comparison with Dizygotic twins (DZ). Although the MZ concordance is not 100% as would be expected for a Mendelian disease, there is a five-fold or greater increased risk of IBD in first-degree relatives with either disease than in the general public (Xavier and Podolsky, 2007), which substantiates the fact that there is a significant genetic contribution towards the development of IBD. Thus, the findings on families (Ohmen et al., 1996) also point to the idea that some genes will be

phenotype-specific, while some overlap in susceptibility loci is to be expected between the two sub-types (Giallourakis et al., 2003; Parkes et al., 2000).

Although both types of IBD involve chronic intestinal inflammation, they differ from each other in several genetic and phenotypic respects, including cytokine profiles and intestinal location of disease manifestation. CD primarily affects the ileal region of the intestine in a transmural (penetrating all layers) fashion. However, CD is also characterized by segmental inflammation (segments of the intestine affected) as well as extra-intestinal manifestation. Unlike CD, the inflammation in UC is uninterrupted and mostly confined to the mucosal layer of either the proximal or distal colon (large intestine, see Figure 1). Table 1 summarizes the overlapping and distinct features between these two sub-types of IBD.

Table 1. Differences and similarities between CD and UC

	Crohn's Disease	Ulcerative Colitis
Prevalence	26-199 per 100 000 (Loftus, 2004)	37-246 per 100 000 (Loftus, 2004)
Disease manifestation	Can affect the entire gastrointestinal tract, from mouth to anus Inflammation is chiefly located in the ileocaecal region. Intestinal inflammation has a discontinuous (segmental) pattern, where affected parts of the gut can be separated by healthy tissue. Can affect all layers of the gut (transmural inflammation)	Inflammation is confined to the colon. Inflammation affects the proximal/distal end of the colon. Intestinal inflammation is continuous, first occurring in the rectum and then extends proximally. Mucosal inflammation only.
Genetic Contribution*	~50%	~10%
Environmental Factors	Smoking, diet	Diet, smoking and appendectomy are protective for UC (Budarf et al., 2009)

* Estimated from Monozygotic/Dizygotic twin concordance rates.

It is understood and widely accepted that the pathogenesis of IBD arises from the complex interplay between the innate and adaptive immunity of a susceptible individual, environmental triggers and the host microbiota (Xavier and Podolsky, 2007). This interplay increases the level of complexity in unraveling the pathogenesis of IBD since the contribution of environmental factors and that of the genetic inheritance of several susceptibility loci and the interaction of these is still unknown (See Figure 2).

Figure 1. The Human gut and structure of the small-intestinal epithelial

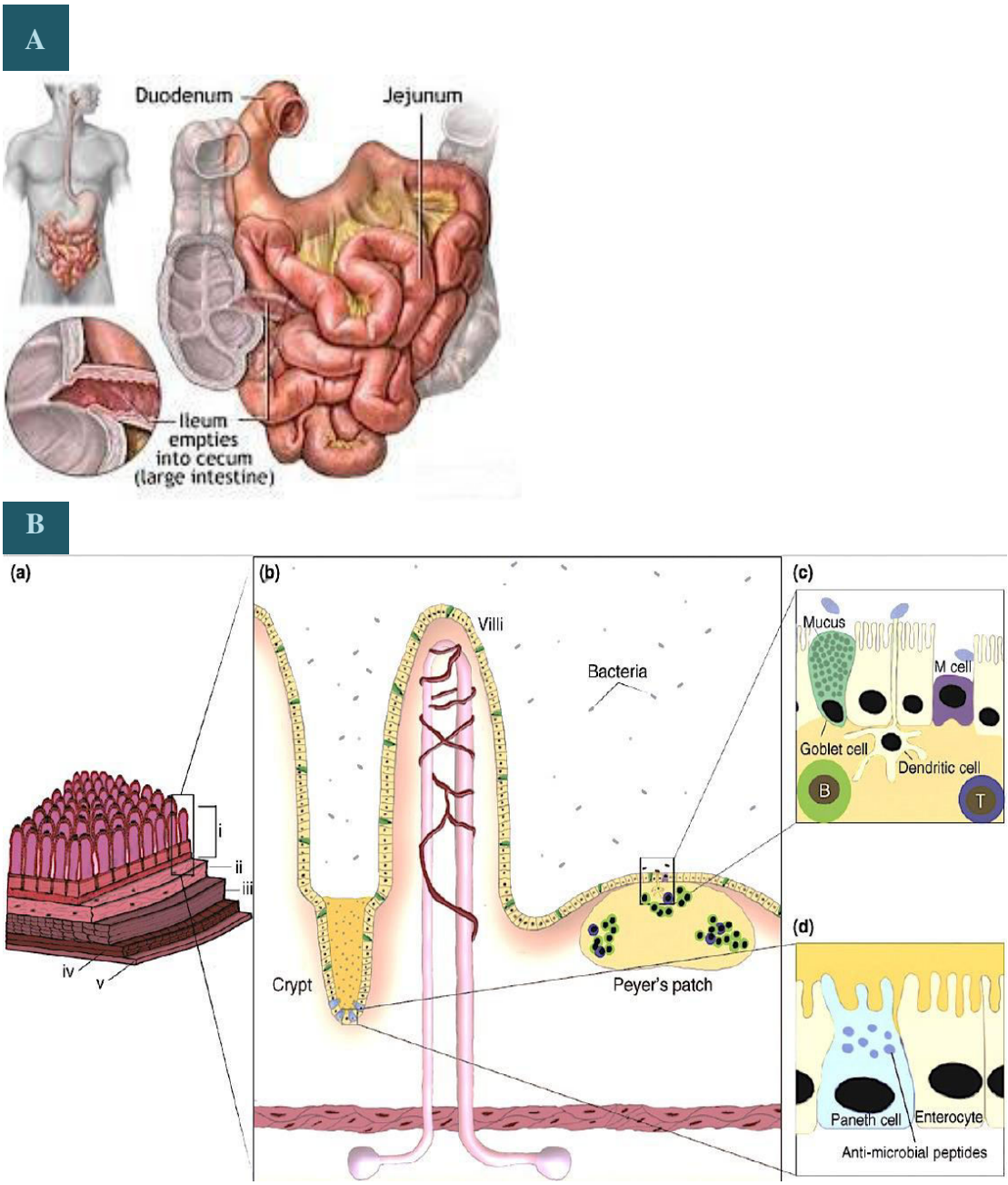


Figure 1. [A] Structure of the human intestine depicting the ileal region where the most common form of CD occurs **[B](a)** The five layers of the gut.**i:** mucosa (the inner-most layer), **ii:** the submucosa, **iii-iv:** underlying muscular layers, **v:** serosa, lining the outer muscular layer. **(b)**Structure of the small intestinal epithelium containing cells important in the gut's biochemical barrier, for example the Goblet cells**(c)** and the Paneth cells **(d)**. (**[A]** *Adapted from* http://www.nlm.nih.gov/medlineplus/ency/imagepages/19221.htm**[B]** *Adapted from Budarf et al, 2009*).

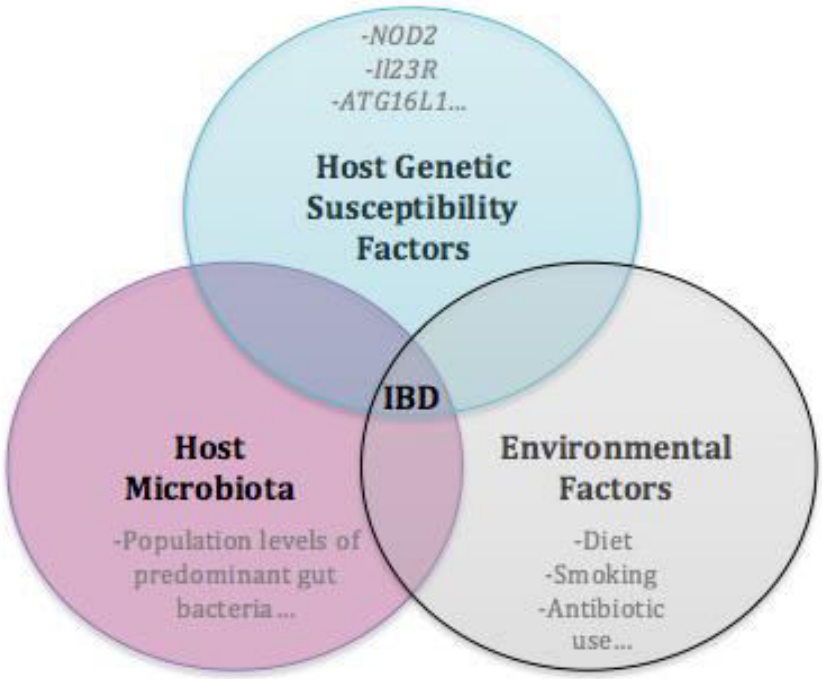


Figure 2. Interaction of the major IBD-contributory factors.
Illustration of the major factors that influence and interact to contribute towards the development of IBD. The individual contribution of each of these components is still unclear, although the environmental component in the development of UC is much stronger than that of CD development.

II. AN OVERVIEW OF THE IMMUNE RESPONSE AND INFLAMMATORY BOWEL DISEASE

The anatomical and physiological properties of the intestinal tract play an important role in the mechanism of IBD. The gut is made up of five different layers, namely the mucosa, consisting of the inner-most layer, which lines the gut lumen, the submucosa, an inner and outer muscular layer and the serosa, which lines the outer muscular layer. (See Figure 1B(a). The epithelial cells of the mucosa are in close contact with the venous system (at their base), as well as being in direct contact with the gut contents at their apical surfaces. The mucosal epithelial layer contains several different cell types (See Figure 1B(b,c,d). The mucus-secreting Goblet cells (See Figure 1B(c) are responsible for forming a protective layer on the luminal surface of the gut while the Paneth cells (See Figure 1B(d), located in the ileal crypts, drive the production of anti-microbial peptides (Wehkamp et al., 2007) and form the biochemical barrier of the gut. The correct function of this dual barrier represents a crucial first defence mechanism in the innate immune system (Budarf et al., 2009), which is the first response upon epithelial penetration of microbes.

If the gut epithelial barrier is penetrated, White Blood Cells (WBCs) are immediately recruited to the epithelial surface where their phagocytic properties come into play. The specialized receptors on WBCs, called Pattern Recognition Receptors (PRRs), are able to recognize the patterns of macromolecules present on the bacterial cell wall and hence initiate phagocytosis (cell-eating), which engulfs and destroys microbes. The recruitment of these cells initiates a cascade involving the production of the necessary cytokines and chemokines as well as triggering the killing of microbes, which will lead to the appropriate signals required to initiate the adaptive immune system (Budarf et al., 2009).

In the situation where the response from the innate immune system fails to contain the microbial invasion, the adaptive immune system is activated, mediated by the specialised B and T Lymphocytes.

III. THE GUT MICROBIOTA AND INFLAMMATORY BOWEL DISEASE

The involvement of bacteria in the pathogenesis of IBD is well established, both through lab based studies of the gut bacteria, which show abnormal bacterial flora (microbiota) in IBD and through association studies on IBD that identified susceptibility genes involved in bacterial recognition. However, the exact role of the microbiota in the pathogenesis of IBD is still unclear. In healthy conditions, the gut microbiota plays a commensal as well as beneficial role in the human gut. The microbiota assists the nutrition of the human host, via the production of short chain fatty acids and vitamins, such as vitamin K, vitamin B12 and folic acid, but also prevents the colonisation of the gut by harmful bacteria (Mai and Draganov, 2009). In contrast to this commensal role, the gut bacteria carry antigens such as lipopolysaccharide and peptidoglycans which can stimulate an immune response. There is a crucial fine balance between host immune tolerance and response and the gut microbiota. Dysbiosis, which is the imbalance of microbial communities inside the digestive tract, of the mucosal and faecal microbiota has been reported for both CD and UC (Cobrin and Abreu, 2005; Targan and Karp, 2005). A loss in microbial diversity is also commonly observed in IBD patients (Ott et al., 2004).

In the case of CD, dysbiosis of the faecal microbiota has also been reported in the unaffected family members of CD patients, when compared to healthy controls in the general population. Joossens *et al* (Joossens et al., 2011) reported that this dysbiosis in the CD patients was mainly characterised by five predominant bacterial species. A decrease in the population levels of *Bifidobacterium adolescentis*, *Dialister invisus*, *Clostidrium* cluster XIVa (uncharacterised species), and *Faecalibacterium prausnitzii* and an increase in levels of *Ruminococcus gnavus* have been identified in CD patients (Joossens et al., 2011). *F. prausnitzii* are bacteria that produce butyrate, which have anti-inflammatory properties, and even lower numbers of *F. prausnitzii* are seen in patients who have had an ileocaecal resection (Joossens et al., 2011). *R. gnavus* do not produce butyrate but are known to degrade gastrointestinal mucin and exhibit high levels of β-glucuronidase (Beaud et al., 2005). This elevated enzymatic activity can lead to the production of compounds in the colon which are toxic and possibly lead to local inflammation (Joossens et al., 2011). This could suggest that increase in the levels of *R. gnavus* in certain regions of the colon and not others could at least in part explain the segmental nature of the inflammation, which is characteristic of CD. In the same study, the unaffected members of the CD families exhibited a decrease in the population levels of *Collinsella aerofaciens*, a member of the *Escherichia coli-Shigella* group and an increase in that of *Ruminococcus torques*. As *R.gnavus*, *R. torques* do not produce butyrate but have gastrointestinal mucin-degradation capacities. These findings suggest that there is a lack of butyrate-producing activity in CD patients, coupled with enhanced mucin-degradation capacity.

IV. GENETIC STUDIES

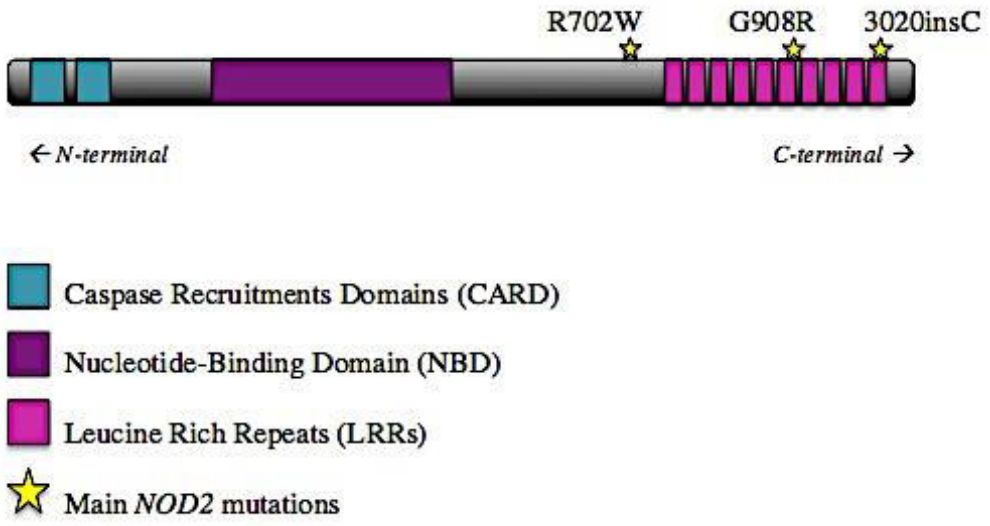
As previously shown in Table 1, CD has a disease prevalence of somewhere between 26 to 199 per 100,000 in populations of European ancestry (Loftus, 2004). Although it is a complex disease there is strong evidence for genetic susceptibility. The estimated sibling recurrence risk ratio (λs) ranges from 15-35 (Schreiber et al., 2005), where λs is defined as the ratio of disease manifestation, given that one sibling is affected, compared with the disease prevalence in the population. However, it was reported earlier (Satsangi, 1994) that the highest λs of 35 was found in a study that only included patients with early onset CD (disease onset prior to the age of 21) and the high values could also be attributable to inclusion of different ethnicities in the studies, in comparison with the general population, in particular Ashkenazi Jewish patients, (Ashkenazi Jewish communities are known to have a high CD prevalence and a tendency towards familial CD). A UK-based study of homogeneous ethnicity and adult-onset disease put the λs estimate closer to 24.7 (Satsangi, 1994). Other studies have shown a high heritability of around 0.50, with the heritability being measured from the difference in correlation between monozygotic and dizygotic twins (Tysk et al., 1988). Because of this high λs and high heritability for a polygenic disease, CD is one of the most widely studied multifactorial diseases.

Early linkage studies, followed by Genome Wide Association Studies (GWAS), were crucial in identifying the first gene involved in IBD. Linkage analysis revealed several different linkage peaks across the genome (Cavanaugh, 2001; Parkes et al., 2000). In particular there was an extensive linkage peak in Non Jewish CD families on Chromosome 16q, at the *IBD1* locus, where the gene *NOD2* is located (Cavanaugh, 2001). Subsequent analyses of this linkage region lead to the first breakthrough, with the discovery that three common *NOD2* (previously known as *CARD15*) variants were associated with CD (Figure 3) (Hugot et al., 2001). The encoded protein, NOD2 is a Nucleotide-binding Oligomerisation Domain protein that has the ability to recognize, intracellularly, microbial cell-wall components, namely Muramyldipeptide (MDP), through its Leucine Rich Repeats (LRRs) (Figue 3). The three *NOD2* mutations identified are located within or close to the LRRs, with the framshift mutation resulting in truncated version of the protein. No mutations within *NOD2* that ablate the production of the entire protein have yet been identified (Eckmann and Karin, 2005). Functional evidence of the three identified *NOD2* mutations, which are illustrated in Figure 3 and in Table 2, will be further discussed in Chapter II.

Table 2. The CD-associated *NOD2* mutations.

rsID	*Change*	*Type*
rs2066844	R702W	Missense
rs2066845	G908R	Missense
rs5743293	3020insC	Frameshift

Figure 3. Representation of the *NOD2* gene and the relative positions of the three common CD-associated *NOD2* mutations (rs2066844- R702W, rs2066845- G9608R and rs5743293- 3020insC).



NOD2 is constitutively expressed in cells of the myeloid lineage including neutrophils and macrophages and its discovery thus provided the first genetic link between CD and an aberrant immune system (Budarf et al., 2009; Eckmann and Karin, 2005; Hugot et al., 2001). However, the strong evidence of linkage on Chromosome 16q could not be fully accounted for by *NOD2* alone since linkage on 16q is still observed in families that do not carry the common *NOD2* mutations. This suggested that there were other, yet unreported, susceptibility genes in the 16q region.

UC is the more prevalent form of IBD, with a disease prevalence of around 37-246 per 100,000 in populations of European ancestry (Loftus, 2004). The estimated siblings recurrence risk ratio λs is approximately 8.5 (Orholm et al., 1991). UC has a lower genetic component than CD, with heritability estimated at about 10% (Thompson and Lees, 2011). Despite the complex nature of UC and the lower genetic contribution, early linkage studies still successfully identified linkage peaks for UC, notably on chromosomes 1, 4, 6, 12, 22 and X, some of which were shared with CD (Hampe et al., 1999; Parkes et al., 2000). Unlike for the case of CD, no genes were pin-pointed and linkage was not observed for UC at the *IBD1* locus (*NOD2* region) on chromosome 16q (Ohmen et al., 1996). This was one of the first indications for the existence of genetic differences between CD and UC, with *NOD2* being a CD-specific factor with a large effect.

Subsequent advances in genomics, including the creation of the International HapMap Project, followed by the advent of high-throughput Single Nucleotide Polymorphism (SNP) genotyping made possible large-scale GWASs, and it is data from these that will be used in my project. HapMap is a multi-national effort that aims to catalogue human genetic variation in order to help scientists identify susceptibility genes predisposing to disease as well as genes involved in individual responses to pharmacological drug metabolism and gene-environment interactions (Risch and Merikangas, 1996).

The term polymorphism is usually applied to variation which is present at a frequency of 1% or more in the general population. Although the data that will be discussed in this thesis will consist of genome-wide SNP data, there are other types of genetic variation present in the human genome. Genetic variation can be subdivided into small-scale sequence variation and large-scale structural variation. The latter include Copy Number Variation (CNV), which involves the loss or gain of a copy of a sequence segment, which may include part or all of a gene, as well as Chromosomal Rearrangements such as translocations and inversions. Small-scale sequence variation includes Variable Number Tandem Repeats (VNTRs), such as Microsatellites, which consist of simple-sequence repeats in the genome. Microsatellites have been extensively used in Linkage studies as they exhibit substantial heterozygosity due to the fact that they are multi-allelic. Base insertions or deletions, also known as Indels, are also a common source of genetic small scale sequence

variation in addition to SNPs. While SNPs (due to allelic imbalance) can be used to detect CNVs genome-wide, CNVs are hard to analyze in genome-wide studies. The same applies to VNTRs, and microsatellites. Small indels are also not included in SNP arrays and all these types of variations may be important particularly in regulatory regions of the human genome.

i. Association Studies

Before the advent of GWAS, SNPs in candidate gene/regions, often but by no means always selected as the result of linkage studies, and most often with a strong hypothesis of putative function, were genotyped and tested for association with the underlying phenotype. Candidate gene association studies that were selected as a result of linkage studies would attempt to narrow down the very large linkage regions, generally spanning more than one megabase in length, in order to ultimately pin-point the causal variant. Indeed, this was the case for the breakthrough with *NOD2* in relation to CD, where the linkage was followed up through fine-mapping by positional cloning and subsequent candidate gene analysis (Hugot et al., 2001). Linkage analysis comes with several advantages, particularly the ability to detect genetic variants of large effect with just a thousand genome-wide genetic markers from a few transmissions within the families studied. Linkage studies are also robust to confounding effects and information can be combined across families even in cases where there are different mutations presenting at the same locus. For these reasons, linkage analysis performs well for

highly penetrant, single–gene disorders but success has been more difficult when linkage analysis was employed for complex traits, which are polygenic conditions thought to be caused by a multitude of variants of low effect-size. Thus, association studies have become the preferred approach in identifying susceptibility loci for complex conditions (Risch and Merikangas, 1996). Although association studies are not robust to effects resulting from population structure, association studies are not limited to families and it is thus easier to recruit a large number of patients and they provide better power for detecting common variants of low effect sizes. However the false-positive rate is higher in association studies than it is for linkage studies.

A GWAS consists of the hypothesis-free analysis of SNPs across the genome to investigate whether the genotype (genotypic test) or allele (allelic test) frequencies for each SNP is significantly different between cases (group affected by the trait) and ethnically matched controls (healthy group). A statistically significant difference provides evidence for association between a particular marker and the trait under study, which requires replication in an independent cohort. SNP genotyping platforms currently used in GWAS can accommodate around one million SNPs, beneficially increasing the SNP resolution across the genome but also giving rise to multiple testing issues due to the individual testing of each SNP. The consequent lack of power means that meta-analyses are now used to improve the number of genetic hits. A meta-analysis combines genotypic or allelic test statistics for the individual SNPs across all the independent GWASs in order to increase

power to detect association by having a much higher sample size. The significant results from a meta-analysis require validation and are followed up in an independent replication sample.

Despite their lack of power, the hypothesis-free nature of GWASs, coupled with the exponential improvement in high-throughput SNP analysis, has nowadays made GWAS the most favourable type of population-based disease-association mapping for complex diseases.

ii. Linkage Disequilibrium

Linkage analysis relies on current recombination, and is thus low resolution due to the limited number of markers that can be informative to obtain linkage phase since there are only a few meiosis within the families studied whereas association relies on the presence and structure of Linkage Disequilibrium (LD), which reflects distant shared ancestry and therefore many recombination events, both current and historic, providing a much higher resolution than linkage. LD is the non-random association of alleles at two closely-linked loci within a random-mating population. Recombination seems to be the major cause of LD breakdown, since LD patterns mirror recombination (Maniatis et al., 2002). However, natural selection, genetic drift, population subdivision and population bottlenecks are all factors that influence LD structure (Kong et al., 2010). Historically recent mutations are likely to be

in high LD with the surrounding SNPs, even in regions of recombination, but are likely to be rare unless selection or demographic processes have increased their frequency.

The extent of LD can be measured using several metrics although the most common include the covariance (D), the ratio of D to its maximum possible absolute value (D'), association (ρ), correlation (r) and correlation coefficient (r^2) (Tapper, 2007) (see Appendix I for formulae). Most of these metrics heavily rely on allele frequencies with the exception of D' and ρ (Morton et al., 2001; VanLiere and Rosenberg, 2008), which are less sensitive to variations in allele frequencies (Morton et al., 2001). r^2 has more power for selecting tagging SNPs than D' but is less good at signaling lack of recombination of low frequency alleles. It gives a value of 1 where the allele frequencies of the SNPs are the same and there is no recombination, but lower values where one of the two SNPs is rare. r^2 has been extensively used in GWA studies, for example in defining LD intervals in meta-analyses.

Each individual has two homologous chromosomes, one inherited from each parent. Therefore, since genotyping platforms are so far unable to assign SNP alleles to their respective homologues, the genotyped SNPs need to be phased. Phasing is the determination of which SNP allele belongs to the maternal or paternal homologue, which is in turn required in order to determine haplotype blocks. Haplotype blocks are genomic regions that

contain markers exhibiting strong LD with one another and thus tend to be inherited together from generation to generation. With the increasing technological advances and decreasing cost of SNP genotyping, haplotype inference and phasing using statistical approaches has become one of the major areas of statistical methodological expansions. Expectation Maximization (EM) is an algorithm that can be used or underlies many statistical approaches used for haplotype inference. EM utilizes an iterative approach in order to infer the haplotype phase, which maximizes the likelihood, or in other words best describes, the data observed. EM works well for a limited number of SNPs, generally ten, but haplotype phase accuracy declines with a greater number of SNPs and the process becomes computationally costly (Browning and Browning, 2011).

In such cases, Coalescent-based models, such as Approximate Coalescent models and Hidden Markov Models (HMM), provide more accurate haplotype phasing as haplotype frequencies are better modeled *a priori* (Browning and Browning, 2011). The Approximate Coalescent Model assumes the ancestral convergence of alleles back in time, for which the parameters are sometimes estimated using a stochastic EM framework to explore all possible solutions or by Markov Chain Monte Carlo (MCMC) methods (Browning and Browning, 2011) .

Although HapMap contains a reference panel of over three million SNPs, the current genotyping platforms used in GWASs only contain a subset of this reference panel. In fact, in some cases the SNPs only partially overlap between different genotyping arrays. Nevertheless, indirect association can still be observed if the typed SNP is in LD (high r^2 value) with the putative causal agent. The power to detect the untyped putative causal agent depends on the allele frequency, and its effect size, as well as the strength of LD with the genotyped SNP. Meta-analyses combine several different GWASs, often carried out using different SNP genotyping arrays, which despite some overlap, by design they contain different SNPs making the combination of test statistics for the same SNPs across the genome difficult among the different studies. Therefore, imputation has been suggested as a way to create a homogenous subset of SNPs to meta-analyse the different independent GWASs. Imputation is an inference method which takes advantage of association between the genotyped SNPs and untyped SNPs. Genotypes of untyped SNPs are inferred by reference to a panel of densely-typed individuals (Guan and Stephens, 2008) (e.g. HapMap database, 1000 Genomes Project). The imputed genotypes are obtained using a Bayesian approach that uses a predicted distribution that is based on haplotype inference from the observed genotypes (Sterne et al., 2009). Imputation has been used universally in GWASs, meta-analyses and other epidemiological studies and introduces small approximations in regions of conserved LD, but can potentially introduce significant bias and inaccuracies in regions of LD breakdown. Imputation algorithms are generally user-defined for each particular dataset in order to carefully model the distribution of each variable

with missing values. Even so, imputation remains an approximation technique. In fact, in a published cardiovascular risk prediction tool that was developed using imputation to handle missing data, it was initially reported that cardiovascular risk was unrelated to cholesterol (Sterne et al., 2009), (Hippisley-Cox et al., 2007). It was however subsequently clarified that when the analysis was carried out using only individuals without missing data, a clear association between cardiovascular risk and cholesterol was observed (Sterne et al., 2009), which highlights the limitations of using imputation for missing data.

In addition to these potential inaccuracies in the meta analyses, the lack of power in detecting genome-wide association could be due to the unrealistic assumption that one of the SNPs on the genotyping platform is either causal or in almost complete LD, as defined by r^2, with the putative causal agent. This would not be the case for rare variants nor in regions of breakdown of LD. In fact it is now widely accepted that the currently used standard methods of analysis are underpowered for detecting rare variants, which could be responsible for substantial allelic heterogeneity in complex diseases (Li and Leal, 2008).

Secondly, another reason for the lack of power comes from the stringent significance thresholds required as a result of multiple testing. Additionally, the P-value for each individual SNP across datasets does not need to reach

nominal statistical significance across all the independent GWASs in order for the P-values in the datasets to be meta-analysed. Another problem of the meta-analysis approach is that it inevitably leads to the pooling of several phenotypic subgroups, as well as different ages of onset (early-onset and adult-onset) as well as ethnicities in order to increase the sample size and hence power to detect association.

iii. Crohn's Disease and Ulcerative Colitis GWASs

The relatively large λs makes CD a good case for genetic studies. In fact, initial studies focused more on CD than UC. One of the first large-scale GWAS was conducted by The Wellcome Trust Case Control Consortium (WTCCC), which analysed ~14,000 cases of seven common diseases, including CD (WTCCC, 2007). The UK WTCCC Phase I CD GWAS analysis identified an initial 9 CD susceptibility loci using ~2,000 cases (WTCCC, 2007). Several other GWASs investigated CD associations in different Caucasian populations, including Belgian (Libioulle et al., 2007), German (Hampe et al., 2007) and USA (McGovern et al., 2010), as well as different ethnicities, namely Jewish and Non-Jewish from North America (Rioux et al., 2007). The largest CD association study to date consisted of a meta-analysis of six of the available independent GWAS datasets, one of which only consisted of early-onset CD patients (Imielinski et al., 2009), combining 6,333 cases and 15,056 controls (Franke et al., 2010). This meta-analysis identified

and confirmed a total of 71 intervals (loci) for CD susceptibility. However, despite the large sample size, only *NOD2* was detected on 16q, despite expectations of other loci on that chromosome. Additionally, in several cases, the LD intervals were quite large (over 1 Megabase) with several suggested genes within them. Despite the international efforts, the confirmed 71 CD loci found up until that time through the several GWASs and meta-analyses were reported to account for only 23.2% (Franke et al., 2010) of the genetic contribution, suggesting that there were many more loci to be found and that further research is required to fully understand the genetics of CD.

Following the WTCCC Phase I studies that encompassed CD, the WTCCC Phase II data generation extended the list of phenotypes under investigation, and now incorporated UC as one of the datasets. The WTCCC Phase II UC data consisted of a larger sample size than the WTCCC CD data, which is more convenient for detecting association in situations like UC, where the genetic contribution is lower than for other complex conditions like CD (Consortium et al., 2009). Several other GWASs from a variety of geographic locations, including Germany, Belgium, Japan and Korea were carried out (Anderson et al., 2011; Asano et al., 2009; Consortium et al., 2009; Ellinghaus et al., 2013; Franke et al., 2008; Franke et al., 2010; Hampe et al., 1999; Haritunians et al., 2010; McGovern et al., 2010; Silverberg et al., 2009; Yang et al., 2013).

More recently, with the aim of expanding the knowledge of the aetiology of IBD, a large-scale genome-wide meta-analysis was carried out, combining data from 15 GWASs of CD and/or UC together with Immunochip genotyping data, where the latter genotyping platform consisted of specific SNPs within or in regulatory regions of immune-related genes. The sample size consisted of more than 75,000 cases and controls together, making this study the largest IBD meta-analysis to date (Jostins et al., 2012) and will be discussed in further detail in Chapter IV. This study identified and replicated 30 CD-specific, 23 UC-specific and 110 UC and CD shared susceptibility loci, resulting in 163 confirmed general IBD susceptibility loci (Jostins et al., 2012). In the European population, the total number of confirmed UC loci now stands at 59, explaining, it is claimed, approximately 16% of the genetic contribution (Jostins et al., 2012).

V. ANOTHER APPROACH- LINKAGE DISEQUILIBRIUM UNIT MAPS

Another approach to association studies is to combine information from many adjacent markers, such as the multi-marker method developed by Maniatis *et al* (Maniatis et al., 2007; Maniatis et al., 2005). The power of this multi-marker method comes partly from the fact that it does not assume the putative causal agent to be in high LD, as measured by r^2, with the genotyped SNP since SNPs within a genomic window are simultaneously analysed within a composite likelihood framework, and partly from the fact that this method incorporates the structure of LD in the human genome by modeling on genetic distances in LD Units (LDUs) in contrast to physical location in kb.

LDU maps represent the genetic distance of SNPs, as opposed to their physical distance. Conventional genetic maps in centi Morgans (cM) are based on recombinants from family data. Such maps can only capture the recently-occuring recombination events since there are only a few generations involved in family data. The main limitation of genetic maps in cM is their relative low resolution since only informative, in other words markers exhibiting heterozygosity in relevant members of the pedigree, can be used for such maps. Microstallite (STR) polymorphisms are ideal, but not frequent enough in the genome. The later genetic map is based on approximately 300,000 informative SNPs on Utah Residents samples with Northern and Western European ancestry from the Centre d'Étude du Polymorphisme

Humain (CEPH) collection, together with families from the deCODE project (Kong et al., 2010). On the other hand, LDU maps provide information from both current and historic recombination events based on the shared ancestry of unrelated individuals and they can be constructed from millions of markers. LDU maps are generally constructed from SNP data using an adaptation of the Malécot model of isolation by distance, which describes the decline of pairwise SNP association as a function of physical distance as shown in Figure 4, with cumulative distances in LDU (Maniatis et al., 2002).

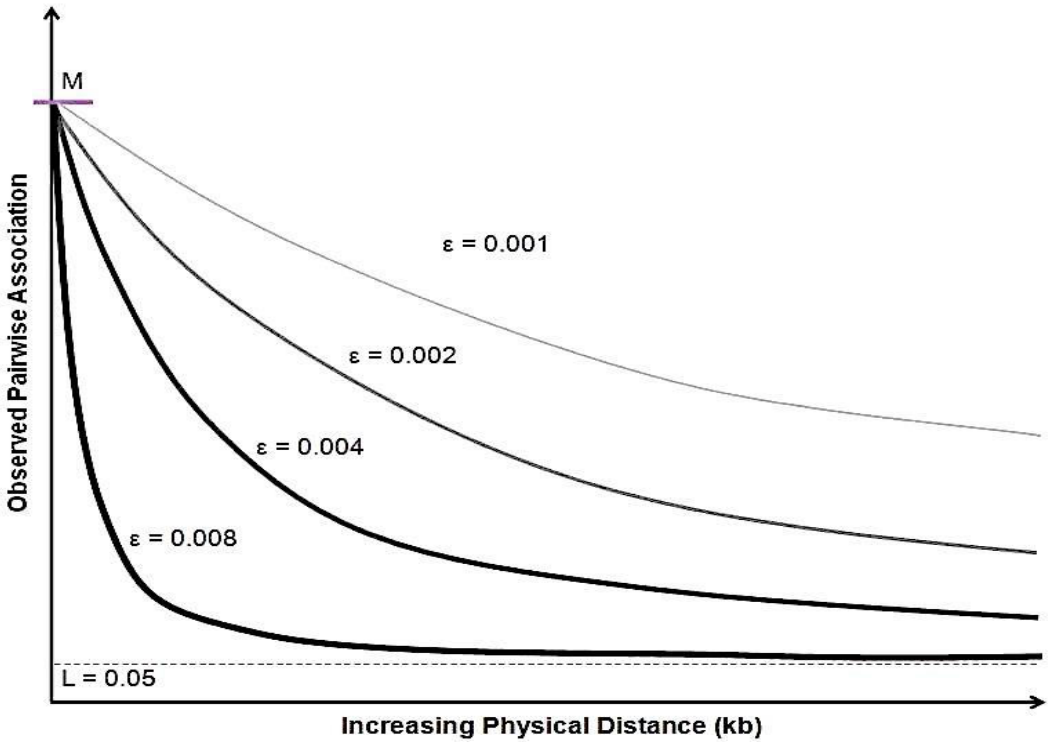


Figure 4. The Relationship between LD and physical distance
Illustrates the best-fit line of various estimates of the observed pairwise SNP associations ($\rho_{(obs)}$. ε is the exponential decline in association ρ as a function of increasing physical distance between pairs of SNPs. The greater the exponential decline ε of pairwise SNP association, the faster the decay of LD (as shown at ε=0.008).

The methodology for constructing LDU maps is given by Maniatis et al (Maniatis et al., 2002) using the LDMAP software (Lau et al., 2007) and is an extension of earlier work (Morton et al., 2007). It is based on the Malécot model that describes the expected decline of association $\rho_{(exp)}$ between any pair of SNPs as:

$$\boldsymbol{\rho_{(exp)} = (1 - L)Me^{-\Sigma\varepsilon_i d_i} + L}$$

where $\rho_{(exp)}$ is the expected association, M is the initial value of LD before decay begins, thus representing the intercept and L is the asymptote. M reflects the phylogenetic origin of haplotypes, where M=1 indicates a unique haplotype and hence a monophyletic origin, and M less than 1 suggests a polyphyletic origin, and thus the presence of more than one ancestral haplotype (Collins and Morton, 1998). L is the asymptote, which corrects for spurious associations at large distances. The exponential decline ε_i represents the decline of association as a function of physical distance (d_i) between a pair of SNPs, with d_i representing the distance in kb between a marker i and a neighbouring marker i+1. The product of $\varepsilon_i d_i$ is analogous and more accurately known than the product of recombination θ and t time in generations and hence $\varepsilon_i d_i$ is a better metric for LD (Tapper, 2007). The observed pairwise SNP-by-SNP associations calculated using the parameter $\rho_{(obs)}$ are used to iteratively estimate, using Maximum Likelihood, the parameters M, ε and L.

To do this, it is necessary first to assign phase, which is done using the Hill algorithm (Hill, 1974), which uses a maximum likelihood approach to infer phase for each pair of SNPs from the frequency of diploid genotypes. The 3-by-3 diplotype table can then be converted into a 2-by-2 table, from which the allele frequencies *Q* and *R* and the covariance *D*, are calculated. An example of the 2-by-2 haplotype table is illustrated in Table 3:

Table 3. 2x2 count pf haplotype frequencies.

		SNP B		
		1	**2**	
SNP	**1**	$a = f_{11}$	$b = f_{12}$	***Q*** *(a+b)*
A	**2**	$c = f_{21}$	$d = f_{22}$	***1-Q*** *(c+d)*
		R *(a+c)*	***1-R*** *(b+d)*	

f is the frequency of the inferred haplotype and (1-*Q*) and (1-*R)* are the allele frequencies.

The covariance *D* can then be calculated as follows:

$$D = f_{11}f_{22} - f_{12}f_{21}$$

The observed association $\rho_{(obs)}$ can then be calculated as follows:

$$\rho_{(obs)} = (ad - bc) \Big/ (a+b)(b+d)$$

which is equivalent to:

$$\rho_{(obs)} = D \Big/ Q(1-R)$$

where *D* is the covariance between the two markers, and *Q* and *R* are the allele frequencies calculated from the 2-by-2 haplotype table illustrated in Table 3. The difference between *D* and *D′* is that for the latter, *D* is normalized by the theoretical maximum and minimum for the observed allele frequencies relative to the value of *D*, as is shown in Appendix 1.

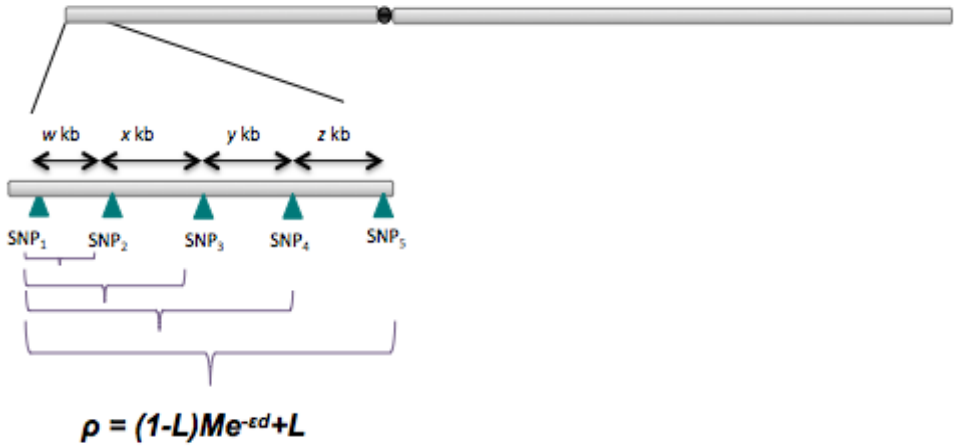


Figure 5. Expected Pairwise Association per SNP interval.
This figure illustrates the first five SNPs of a chromosome, with the first SNP interval composed of pairwise SNP association between SNP_1-SNP_2, SNP_1-SNP_3, SNP_1-SNP_4, SNP_1-SNP_5, which are used to estimate values of ε.

Figure 5 illustrates the first five SNPs on a chromosome, where the first SNP interval is composed of pairwise associations between SNP_1-SNP_2, SNP_1-SNP_3, SNP_1-SNP_4, SNP_1-SNP_5 …SNP_1-SNP_n. For each *i*th SNP interval, the Malécot model

$$\boldsymbol{\rho_{(exp)} = (1 - L)Me^{-\Sigma\varepsilon_i d_i} + L}$$

is subsequently used to iteratively estimate, using a maximum likelihood approach, the values of *M*, ε and *L,* by using all the observed pairwise association for the SNPs covering that specific interval, and their corresponding distances d_i in kb, weighted by the information $K\rho$ for each interval (Tapper, 2007). This information is now used to construct LDU maps where:

$$LDU = \varepsilon_i d_i$$

for every SNP interval.

These LDU maps can describe the underlying structure of LD (Maniatis et al., 2002) and are visualized by plotting marker locations in LDU against distances in kb, which demonstrate the non-linear relationship between physical distance and LD. These metric genetic maps are analogous to linkage maps in cM (Tapper et al., 2005). Blocks of conserved LD are areas of reduced haplotype diversity while steps represent LD breakdown most likely due to recombination as cross-over profiles agree with LDU patterns (Webb et al., 2008). Although patterns of LD are mainly a consequence of recombination (Jeffreys and Neumann, 2009), other factors, such as selection and genetic drift, as well as stochastic effects are important (Slatkin, 2008). HapMap Phase II data is used to obtain genetic distances in LDU since their estimation is based on a higher resolution as a result of a denser SNP coverage as opposed to HapMap III.

Subsequent to the construction of LDU maps, for each analytical window the same Malécot model is used to model the decline of affection-status-by-SNP association as a function of genetic distance in order to estimate the location of the putative causal agent. This will be discussed in more detail in Chapter II. One of the major advantages of using composite likelihood over

single-SNP testing is a greatly reduced number of tests and hence a less stringent multiple-testing correction. This method therefore addresses the current major problem of lack of power associated with single SNP testing and avoids the need for imputing as a result of analysing SNPs simultaneously in analytical windows. Other details are given in Maniatis *et al* (Maniatis et al., 2007), which provides evidence of the power and resolution of this approach over single SNPs.

VI. GENERAL AIMS OF STUDY

The overall aim of this study was to make progress on elucidating the complex genetics of IBD using the multi-marker LD mapping approach described in the previous section (Maniatis et al., 2007). The first step in this study was to re-analyse Chromosome 16q in relation to CD, on the publically available WTCCC CD data and the North American National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK) IBD Genetics Consortium CD data (IBDGC) as a replication sample, with the aim of identifying novel susceptibility genes on Chromosome 16q, other than *NOD2*. The goal of this part was to shed light on the possibility of genetic heterogeneity underlying the region of significant linkage on 16q in patients who do not carry the common *NOD2* mutations.

The second step was to apply the same multi-marker approach to the whole genome using the same data (WTCCC and NIDDK) for CD. The first aim was to identify whether the 71 previously reported CD loci based on the meta-analysis of six GWASs, could be identified using only the WTCCC and NIDDK datasets. The second was to see whether this method would identify novel susceptibility loci for CD, since the 71 loci did not appear to account for the total genetic contribution.

Subsequently, the attention was shifted towards UC. The same method was applied to the WTCCC Phase II UC data and two NIDDK UC datasets as replication samples, in order to identify novel susceptibility loci for UC.

CHAPTER II. DISSECTING CHROMOSOME 16: GENETIC AND PHENOTYPIC STRATIFICATION PROVIDES INSIGHT INTO CROHN'S DISEASE.

I. INTRODUCTION

One of the major breakthroughs in CD genetics came in 2001 when three functional mutations within *NOD2* were identified, as a culmination of the previous work using family and association studies. *NOD2* encodes a protein that is responsible for recognizing pathogens through its Leucine Rich Repeats (LRRs) domain and hence this discovery provides a direct link to a genetically altered host immune system (Budarf et al., 2009; Hugot et al., 2001).

It was however subsequently reported that there is a lack of direct relationship between the prevalence of these three mutations (rs2066844, rs2066845, rs5743293, see Table 2 and Figure 3 in Chapter I) in the general population and CD frequency (Hugot et al., 2007). Indeed, *NOD2* mutations have not been found in all CD patients (Lesage et al., 2002), although this is not surprising in complex inheritance. It is also important to note that the

NOD2 mutations are also relatively frequent in healthy individuals in Central Europe, and the average frequency of these mutations in Caucasians is 4.3% for R702W (rs2066844), 1.2% for G908R (rs2066845) and 2.3% for the L1007fsinsC frameshift mutation (rs5743293) (Hugot et al., 2007). In addition, the linkage peak originally identified on chromosome 16 persisted even after the patients that were found to have the *NOD2* mutations were removed from the data set, however in this case, the linkage peak shifted some 25cM downstream of the *NOD2* gene (van Heel et al., 2003), suggesting the presence of other chromosome 16q susceptibility genes. Despite the shift in the linkage peak, no other signals on 16q have been found so far in published GWA studies. To add to the complexity, *NOD2* mutations are associated with ileal CD but not with perianal or colonic disease (Ahmad et al., 2002), further suggesting the involvement of both genetic and phenotypic heterogeneity in relation to CD.

At the outset of this thesis work, the IBD Genetics Consortium had made significant progress in mapping genes involved in CD, and combined GWASs identified an initial 32 susceptibility loci (Barrett et al., 2008), which included *NOD2* on chromosome 16q, followed by an additional 39 (Franke et al., 2010), making a total of 71. Several of these loci are implicated in other diseases involving inflammatory and immune dysregulation (thus, *PTGER4* on Chromosome 5p, is common to CD, UC (Anderson et al., 2011) and multiple sclerosis). However, this study reported that the 71 loci account for only

23.2% (Franke et al., 2010) of the genetic risk a problem of so-called “missing heritability”.

In this chapter I describe how the GWA database was revisited using the multi-marker LD mapping approach mentioned in the Introduction (Maniatis et al., 2007). This first analysis focused on the previously identified *IBD1* linkage region on Chromosome 16q (Cavanaugh, 2001), since the strong evidence for linkage in this region cannot be fully accounted for by *NOD2* alone. Indeed, using this LDU multi-marker approach, several distinct association signals of high significance were identified on chromosome 16q alone. For this first step, I focused on three of these signals to study in detail. Each of the three signals was found in the WTCCC data and, quite independently, in the NIDDK dataset, which also contains information on Jewish ancestry as well as additional sub-phenotype data.

The work in this chapter has already been published (Elding et al., 2011) and contributions made by others are mentioned in the subsequent text.

II. METHODS

i. Subjects and Methods

2,009 cases of Crohn's Disease and 3004 combined controls from the GWA scan of the UK WTCCC study (WTCCC, 2007) based on the Affymetrix 500K array, which contains approximately 500,000 SNPs, were analysed in this study. Half of the ~ 3,000 nationally-ascertained controls came from the 1958 British Birth Cohort collection and the remainder from the UK Blood Services Collection of Common Controls (WTCCC, 2007).

For this study, the quality control filtering was conducted by Winston Lau and the procedure that was followed is described in detail in the Methods section of Chapter IV. After quality control, a total of 1698 cases and 2948 combined controls were retained for analysis. The cases were confirmed patients of any subtype of CD using endoscopic, radiological procedures and histopathological criteria. The cases were not specifically enriched for early age of onset or family history and they were from a variety of IBD clinics. For the replication study, following quality control, 813 North American patients with CD and 947 controls made available by the NIDDK IBD Genetics Consortium (IBDGC) were analysed. The NIDDK IBDGC GWA scan was

based on the Illumina HumanHap300 array (Rioux et al., 2007), consisting of approximately 320,000 SNPs, and has a smaller sample size and SNP set that only partially overlaps with the WTCCC SNP array. The WTCCC reports fewer phenotypic data than the NIDDK study does, and the latter includes information on the involvement of other intestinal locations.

ii. Genetic analysis

A high density LDU map for the whole of chromosome 16 had been previously constructed using the CEU (Utah residents with Northern and Western European ancestry) PHASE II data from the HapMap Project (Lau et al., 2007). See Chapter I for details about the construction of LDU maps.

The entire 16q was divided into non-overlapping windows based on the LDU map, with a minimum length of 10 LDU and a minimum number of 30 SNPs for each window (See Figure 6).

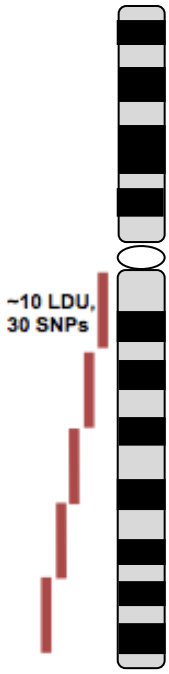


Figure 6. Partitioning the Chromosome.
The chromosome arm is partitioned into non-overlapping 10 LDU windows and a minimum of 30 SNPs in each 10 LDU window.

This resulted in 98 windows for 16q for both GWA datasets. For each window the LDU mapping method previously published by Maniatis *et al* (Maniatis et al., 2007) was used, which will be described in detail in this section. This method returns a P-value and an estimate of the causal location ($\hat{S}$) using genetic locations derived from a high–resolution linkage disequilibrium (LD) map (Maniatis et al., 2002). The association mapping approach that was applied for each window was based on the same Malécot model but ρ was

replaced with *z,* which reflects the association between the affection status and every SNP.

Table 4. A 3x2 table of Affection Status-by-genotype, coverted into a 2x2 table of Affection Status-by-SNP alleles.

Affection Status	***Genotypes***		
	AA	**Aa**	**aa**
Affected **1**	n_{AA}	n_{Aa}	n_{aa}
Unaffected **0**	n_{AA}	n_{Aa}	n_{aa}
Affection Status	***Alleles***		
	A	**a**	**Total**
Affected **Count**	**a**= 2(n_{AA}) +n_{Aa}	**b**= 2(n_{aa}) + n_{Aa}	***f*** a+b
Unaffected **Count**	**c**= 2(n_{AA}) +n_{Aa}	**d**= 2(n_{aa}) + n_{Aa}	***1-f*** c+d
Total	***R***	***1-R***	**1**
Frequencies			
Count	a+c	b+d	n=a+b+c+d

f is the frequency of the affected individuals and *R* is the allele frequency.

The observed association $z_{(obs)}$ can then be calculated as follows (see Table 4):

$$z_{(obs)} = {(ad - bc)}/{(a+b)(b+d)}$$

Which is equivalent to:

$$z_{(obs)} = {D}/{f(1-R)}$$

where *D* is the covariance between the Affection Status and the marker alleles and *f* is the frequency of affected individuals in the population and *R* is the allele frequency.

The objective of this method is to estimate $\hat{S}$, which is the location of the causal variant on the HapMap LDU map. Therefore, the distance *d* in the model is replaced with $|(S_i-\hat{S})|$ where S_i is the LDU location of the *i*th marker within the window and $\hat{S}$ is the location estimate.

$$z_{(exp)} = (1 - L)Me^{-\varepsilon|S_i-\hat{S}|} + L$$

Each window was tested for association with CD by using composite likelihood (Λ) that combines information from all single-SNP tests within each window, and therefore avoids undue multiple testing correction. Λ is estimated as:

$$\Lambda = \sum k_i[z_{(obs)_i} - z_{(exp)_i}]^2$$

where $z_{(obs_i)}$ is the observed association between the affection status and every *i*th SNP and $z_{(\exp_i)}$ is the expected association from the model, weighted for nominal information K_i.

The significance of association for each window was obtained by calculating an *F*-statistic from the comparison between Λ from the null model, assuming no association, and Λ from the alternative, where the $\hat{S}$ is iteratively estimated. Since each window will contain a different number of SNPs, the P-value takes into account the different degrees of freedom for each window. Therefore, for convenience, the *F*-statistic is converted to a χ^2 with 1 degree of freedom (Maniatis et al., 2007).

A Bonferroni threshold was calculated in order to account for multiple testing, using the following formula:

$$\boldsymbol{P - value\ threshold = \frac{\alpha}{Number\ of\ tests}}$$

where α is the commonly accepted probability of a 5% false positive rate. The Bonferroni P-value threshold used was 5×10^{-4}. This P-value threshold corresponds to the 98 analytical windows on chromosome 16q that were tested for this analysis.

The 95% confidence interval (CI) for the estimated location $\hat{S}$ was obtained as:

$$\boldsymbol{\hat{S} \pm t(SE)}$$

where *t* is the tabulated value of Student's *t*-test and *SE* is the Standard Error of the parameter $\hat{S}$. The predicted estimates of $\hat{S}$ and 95% CIs are obtained by fitting the LDU genetic distances since this approach increases the power of association (Maniatis et al., 2007). For convenience, these estimates are converted back to kb (NCBI build 37) by linear interpolation of the two flanking SNPs in HapMap. Therefore the 95% CIs measured in kb cannot be symmetrical because of the 'block-step' structure of the human genome. When $\hat{S}$ is in an LD block (horizontal line) then all markers within that block have the same LDU location. In such cases, the midpoint of that block was taken as an estimate of $\hat{S}$ in kb.

The significance around the $\hat{S}$ region was additionally analysed by calculating a curvature that reflects the likelihood surface. This curvature is based on χ^2_1 estimates at fine resolution. Each χ^2_1 is estimated in the same way as above but the model that estimates $\hat{S}$ is compared against a model, whereby $\hat{S}$ is fixed incrementally. These χ^2_1 values are converted to *Z* estimates using the following formula:

$$\boldsymbol{Z = \frac{\chi_1^2}{-2ln10}}$$

Therefore, these logs of the likelihood odds ratios (LOD) provide a measure of confidence. All *Z* estimates are plotted against the kb map in order to visualise the locations in the map that deviate from the maximum likelihood

estimation (Ŝ), as shown in Figure 7B and Figure 7C. The smaller the Z estimate, the greater the maximum likelihood.

iii. Genetic and Phenotypic stratification

In order to stratify the data with respect to carrier status for the most common *NOD2* mutations, in the absence of actual data on those mutations, linked SNPs were used. rs2076756 was used in the WTCCC dataset (G being the minor allele with a frequency of 0.24, in the controls, and 0.32 in the cases) and rs5743289 for the NIDDK dataset (T being the minor allele frequency of 0.17 in controls and 0.24 in cases). These SNPs, which are approximately a 100bp apart, are in a region of conserved LD that contains the three *NOD2* mutations (see LDU map in Fig 6A; and as previously reported for rs2076756).

The WTCCC dataset contains patients of any subtype of CD but with no sub-classification available in the database. The NIDDK database on the other hand, contains additional information on disease, in particular that of possible extra-ileal intestinal involvement and also classifies the patients and controls according to ancestry (Jewish/Non-Jewish). Since there was prior expectation of genetic differentiation across these categories we exploited this

extra information to stratify the analyses of the *CDH3, CDH1* and *IRF8* regions.

III. RESULTS

Initially, using the WTCCC data for chromosome 16q, three major signals were identified, in the regions of *NOD2/CYLD*, *CDH3/CDH1* and *IRF8* (Tables 4, 6 and 7). The NIDDK GWAS data was then used to replicate this study.

I. *NOD2* and *CYLD* region

Table 5 and Figure 7A show the significance and estimated locations ($\hat{S}$) for the genomic region that harbours *NOD2* and *CYLD* for both datasets. Two separate signals of association mapped to different locations in this region. Analysis of a window that included marker information from *NOD2* but not *CYLD* yielded a highly significant association with CD for the WTCCC dataset, which was replicated using the NIDDK dataset. The estimated ($\hat{S}$) location was identical for both datasets (50,749.2 kb, 95% CI for WTCCC: 50,707-50,766 kb and 95% CI for NIDDK: 50,708-50,839 kb), which is within an LDU block that spans 16 kb. The block includes exons 4 and 8, which harbour two of the most frequent functional mutations (rs2066844- R702W and rs2066845- G908R, respectively) within the *NOD2* gene. The third most frequent mutation (rs5743293- L1007fsinsC) is on a neighbouring LD block with a very slightly different LDU location (see LDU block, Figure 7A).

The analysis of the WTCCC data for the adjacent window containing *CYLD* yielded an estimated location 11 kb ($\hat{S}$: 50,847.3 kb, 95% CI: 50,839-50,851 kb) downstream of the gene (Figure 7A) and approximately 98 kb downstream of the *NOD2* signal. This signal was replicated with the NIDDK data with an estimated location 234 base pairs downstream from the WTCCC $\hat{S}$ ($\hat{S}$: 50,846.5 kb, 95% CI: 50,846-50,847 kb). For the *CYLD* window, the 95% CI was smaller than that obtained for the *NOD2* window because there is less LD in the region.

Analysis of a window that included marker information from both *NOD2* and *CYLD* (*NOD2&CYLD* window, see Table 5) yielded an estimated location $\hat{S}$ between *NOD2* and *CYLD* for the WTCCC data, and the likelihood surface incorporated both genes (Figure 7A, 7B and 7C). The analysis of the same *NOD2&CYLD* window using the NIDDK data produced a very similar likelihood surface that incorporated both *NOD2* and *CYLD* although the estimated location pointed within *NOD2*, very close to the three *NOD2* mutations, as for the window that contained marker information for NOD2 alone (Figure 7A, 7B and 7C).

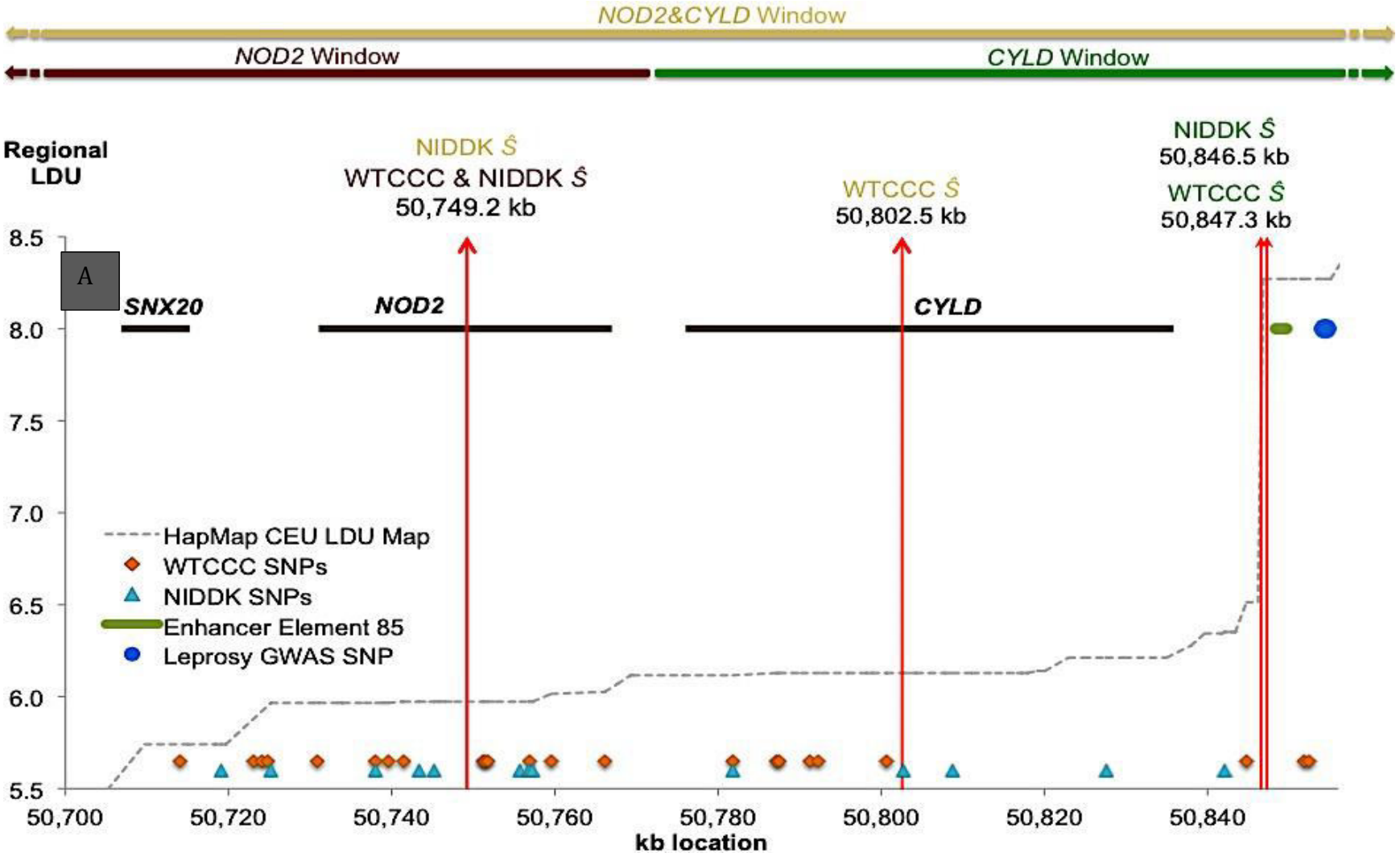
Table 5. Association statistics and estimated location of the causal variation for three different windows covering *NOD2* alone (50,731-50,767 kb), *CYLD* alone (50,776-50,836 kb) and *NOD2* and *CYLD* combined in relation to the locations one the human genome sequence.

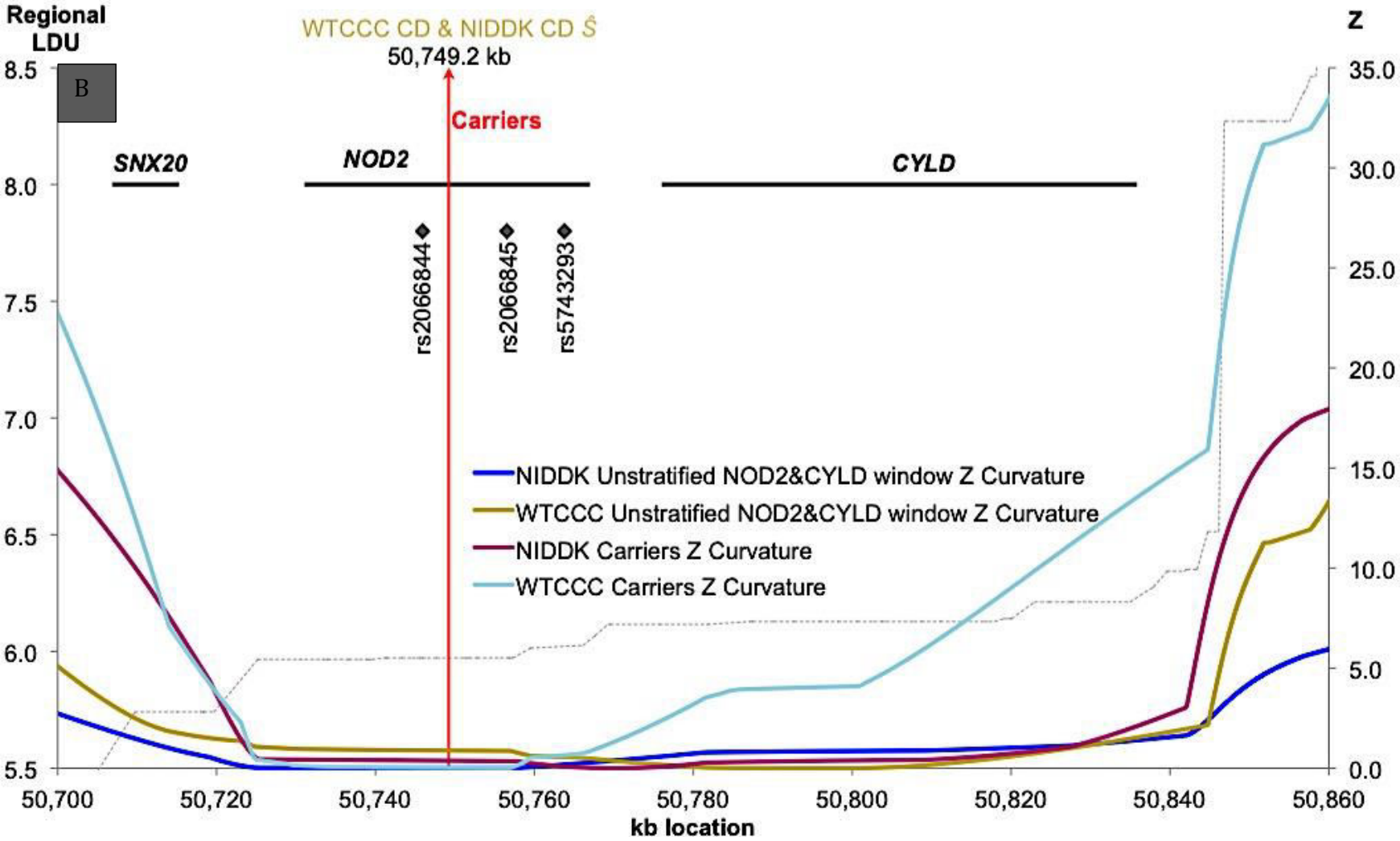
Data	Window[+]	Cases	χ^2_1[†]	P-value	Estimated Location (kb)	95% Confidence Interval CI (kb)
WTCCC	*NOD2&CYLD*	1698	79.7	4×10^{-19}	50,802.5	50,701-50,846
NIDDK	*NOD2&CYLD*	813	31.9	2×10^{-08}	50,749.2	50,708-50,839
WTCCC	*NOD2*	1698	62.6	3×10^{-15}	50,749.2	50,707-50,766
NIDDK	*NOD2*	813	37.1	1×10^{-9}	50,749.2	50,708-50,839
WTCCC	*CYLD*	1698	54.4	2×10^{-13}	50,847.3	50,839-50,851
NIDDK	*CYLD*	813	12.5	4×10^{-4}	50,846.5	50,846-50,847

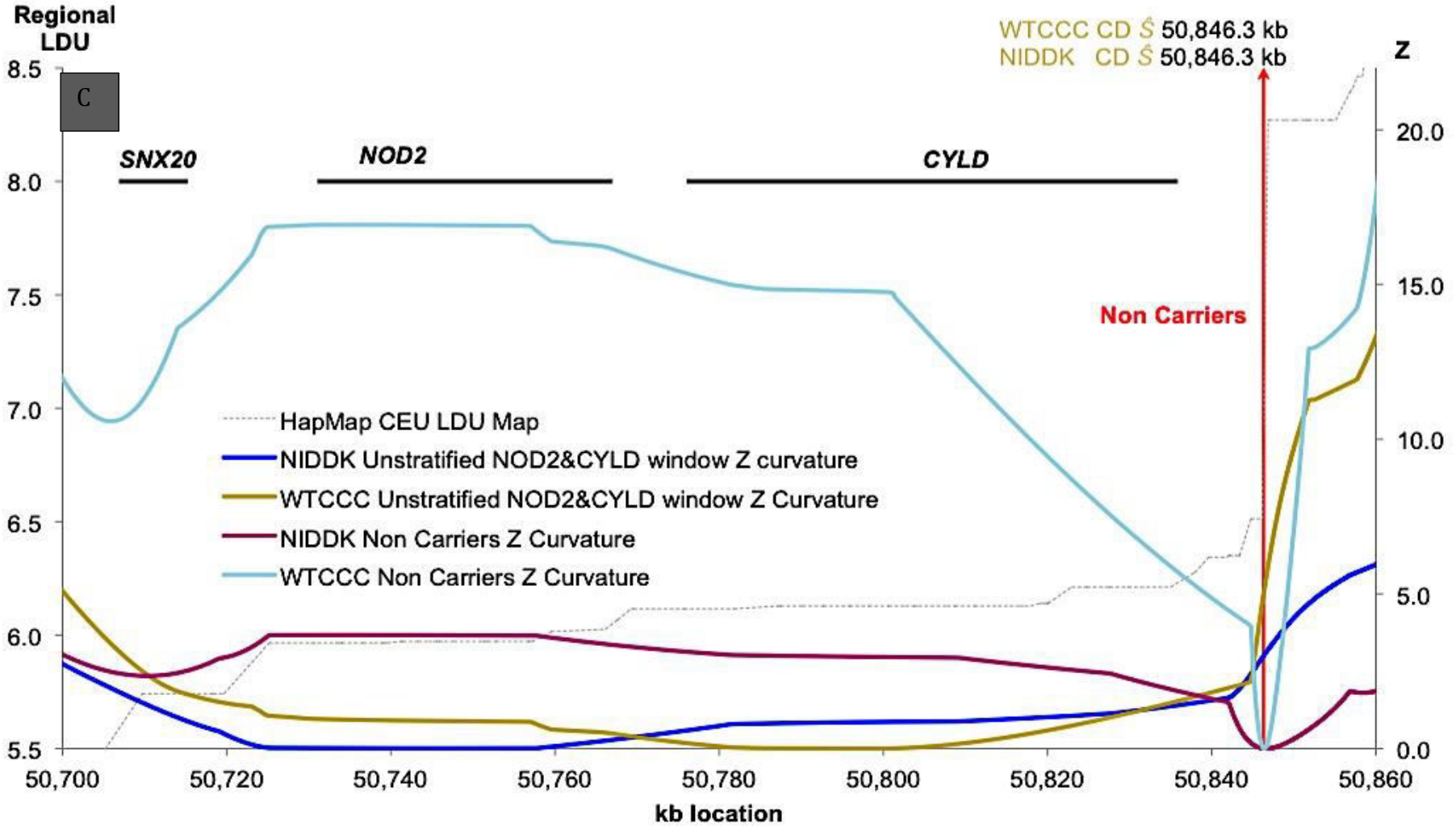
[+]Window with marker information covering *NOD2* or the adjacent window covering *CYLD.*[†]A window χ^2 using composite likelihood.

Figure 7A shows the estimates of $\hat{S}$ for the two different *NOD2/CYLD* windows and datasets. For both datasets, the *CYLD* window was significantly associated with CD (2×10^{-13} and 4×10^{-4} for WTCCC and NIDDK respectively, Table 5).

Figure 7. LDU mapping of the associations detected in the *NOD2*, *CYLD* and combined windows.




Regional LDU
Z
WTCCC CD & NIDDK CD Ś
50,749.2 kb
B
Carriers
SNX20
NOD2
CYLD
rs2066844
rs2066845
rs5743293
NIDDK Unstratified NOD2&CYLD window Z Curvature
WTCCC Unstratified NOD2&CYLD window Z Curvature
NIDDK Carriers Z Curvature
WTCCC Carriers Z Curvature
8.5
8.0
7.5
7.0
6.5
6.0
5.5
35.0
30.0
25.0
20.0
15.0
10.0
5.0
0.0
50,700
50,720
50,740
50,760
50,780
50,800
50,820
50,840
50,860
kb location


Regional LDU
C
WTCCC CD Ś 50,846.3 kb
NIDDK CD Ś 50,846.3 kb
Z
SNX20
NOD2
CYLD
Non Carriers
HapMap CEU LDU Map
NIDDK Unstratified NOD2&CYLD window Z curvature
WTCCC Unstratified NOD2&CYLD window Z Curvature
NIDDK Non Carriers Z Curvature
WTCCC Non Carriers Z Curvature
8.5
8.0
7.5
7.0
6.5
6.0
5.5
20.0
15.0
10.0
5.0
0.0
50,700
50,720
50,740
50,760
50,780
50,800
50,820
50,840
50,860
kb location

Gold font relates to analysis results for the combined *NOD2&CYLD* window. Maroon font illustrates results for the *NOD2* window alone whereas green font represents results for the *CYLD* window alone. [A] Analysis of all patients with SNP data grouped in two separate windows covering either *NOD2* (*NOD2* window) or *CYLD* (*CYLD* window), and analysis of all patients with SNP data grouped in a window combining marker information for both *NOD2* and *CYLD* (*NOD2&CYLD* window). The vertical lines are the estimated locations ($\hat{S}$) for the *NOD2* window, the *CYLD* window and the *NOD2&CYLD* window**.** The grey dashed line shows the LDU map constructed from the HapMap CEU data. **[B] stratified data enriched for carriers of the most frequent *NOD2* mutations.** rs2066844, rs2066845 and rs5743293 are the three major *NOD2* mutations. The dark blue line represents the likelihood surface for the NIDDK *NOD2&CYLD* window analysis of the unstratified data. The gold line represents the likelihood surface for the WTCCC *NOD2&CYLD* window analysis of the unstratified data. The magenta line represents the likelihood surface of the *NOD2&CYLD* window analysis of the NIDDK patients who carry the *NOD2* mutations. The light blue line represents the likelihood surface of the *NOD2&CYLD* window analysis of the WTCCC patients who carry the *NOD2* mutations. **[C] stratified data enriched for non-carriers for the most frequent *NOD2* mutations.** The dark blue line represents the likelihood surface for the NIDDK *NOD2&CYLD* window analysis of the unstratified data. The gold line represents the likelihood surface for the WTCCC *NOD2&CYLD* window analysis of the unstratified data. The magenta line represents the likelihood surface of the *NOD2&CYLD* window analysis of the NIDDK patients who do not carry the *NOD2* mutations. The light blue line represents the likelihood surface of the *NOD2&CYLD* window analysis of the WTCCC patients who do not carry the *NOD2* mutations.

II. Genetic Heterogeneity within the *NOD2* and *CYLD* region

These results showed that the *NOD2* region is more complex than previously thought. The two different *Ŝ* locations within the *NOD2*/*CYLD* region are an indication of the existence of different risk alleles in different patients. The three frequent causal *NOD2* mutations (Hampe et al., 2001; Hugot et al., 2001; Hugot et al., 2007) were not included in the genotyping platforms for both WTCCC and NIDDK and hence the carriers of the mutations could not be identified. However, the data included known SNPs that are in strong LD with these mutations (i.e.: form part of the background haplotype on which the functionally significant mutations arose). The relationship between these two genes was therefore further investigated by analysing a window that included marker information from both *NOD2* and *CYLD*, but stratifying the data according to specific SNPs. The stratification separated, as far as possible, (in the absence of data on the causative *NOD2* mutations), patients without the main functional *NOD2* mutations from those with greatest likelihood of carrying them.

The group that includes all the carriers of the disease-associated allele [heterozygous and homozygous for the minor allele, i.e. carriers with AG/GG or CT/TT genotypes for the rs2076756 (WTCCC) and rs5743289 (NIDDK) SNPs respectively] unsurprisingly yielded much higher significance levels than before, even though the number of patients was much smaller than that of the full dataset (Table 6). This confirms the notion that these two groups of patients included the majority of cases with the functional *NOD2* mutations. In addition, the *Ŝ* was

within *NOD2* despite analysing both genes in the same window (*NOD2&CYLD* window, Table 6, Figure 7B). The position of $\hat{S}$ (50,749.2 kb) was exactly the same as for the *NOD2* window in Figure 7A using the unstratified data.

The analysis of the non-carrier cases for the WTCCC (AA for rs2076756) and NIDDK (CC rs5743289), produced essentially identical results to each other, both, pointing towards a location approximately 11 kb downstream of *CYLD* (Table 6, Figure 7C). This genetic stratification reveals heterogeneity among patients with CD and suggests that a second locus in the vicinity of *CYLD* plays a larger role in patients who do not carry *NOD2* mutations.

Table 6. Association statistics and estiamted location of the causal variation for a window covering *NOD2* (50,50,731-50,767 kb) and *CYLD* (50,776-50,836 kb) in relation to the locations on the human genome sequence.

Data	Window[+]	Cases	χ^2_1[†]	P-value	Estimated Location (kb)	Signal
WTCCC						
AA[*]	*NOD2* &*CYLD*	805	46.0	1×10^{-11}	50,846.3	*CYLD*
AG[*]	*NOD2* &*CYLD*	665	124.6	6×10^{-29}	50,749.2	*NOD2*
GG[*]	*NOD2* &*CYLD*	199	82.8	9×10^{-20}	50,749.2	*NOD2*
NIDDK						
CC[*]	*NOD2* &*CYLD*	482	13.5	2×10^{-4}	50,846.3	*CYLD*
CT[*]	*NOD2* &*CYLD*	266	60.5	7×10^{-15}	50,749.2	*NOD2*
TT[*]	*NOD2* &*CYLD*	60	103.3	3×10^{-24}	50,749.2	*NOD2*

*Data stratified based on the AA, AG, GG genotypes for the rs2076756 SNP from the WTCCC data (50,756.9 kb) and CC, CT, TT genotypes for the rs5743289 SNP from the NIDDK data (50,756.8 kb). [+]Window with marker information from both genes.[†]A window χ^2 using composite likelihood.

III. *CDH3*/*CDH1* and *IRF8*

Table 7 shows the significance and estimated locations for the *CDH3/CDH1* and *IRF8* windows. The original analysis of the WTCCC data had shown that both windows were significantly associated with CD (1×10^{-8} and 6×10^{-9}, respectively). Neither of these signals were however initially replicated with the NIDDK GWA scan of the full data-set but showed significant association when the windows were analysed based on a subset of the data, using phenotypic

information provided by the NIDDK IBDGC. This subset included patients with ileal CD who had involvement of at least one extra-ileal intestinal location, *i.e.* jejunal, colorectal or perianal. The signal near *CDH3/CDH1* was replicated using this subset of the NIDDK data despite the much smaller number of cases (Table 7). For both GWA scans, the estimated location $\hat{S}$ for the former window was between *CDH3* and *CDH1,* within an LDU block that spans 65 kb (Figure 8). The causal locus could be anywhere within that block, which encompasses the 3' region of *CDH3* through to *CDH1* intron 2 (Figure 8), but it also includes a functional promoter SNP rs16260 (Li et al., 2000) (Figure 8) previously associated with Irritable Bowel Syndrome (IBS).

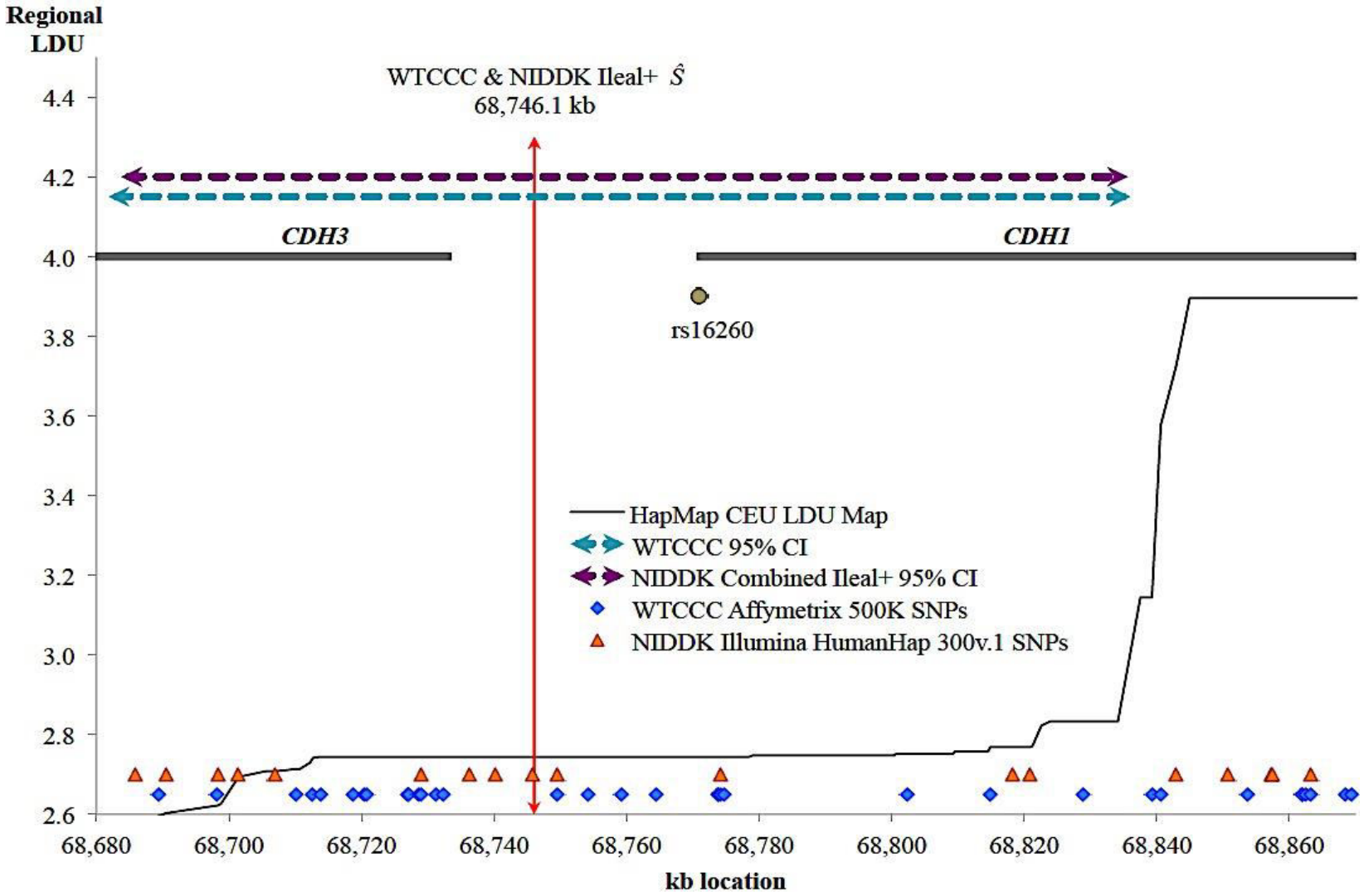

Regional LDU
WTCCC & NIDDK Ileal+ $\hat{S}$
68,746.1 kb
CDH3
CDH1
rs16260
HapMap CEU LDU Map
WTCCC 95% CI
NIDDK Combined Ileal+ 95% CI
WTCCC Affymetrix 500K SNPs
NIDDK Illumina HumanHap 300v.1 SNPs
4.4
4.2
4.0
3.8
3.6
3.4
3.2
3.0
2.8
2.6
68,680
68,700
68,720
68,740
68,760
68,780
68,800
68,820
68,840
68,860
kb location

Figure 8. Localisation within the *CDH3/CDH1* region.
The LD map of the region is illustrated by the black line, which is obtained by plotting HapMap LDUs (Y axis) against kb (X axis). The red vertical arrow is the estimated location $\hat{S}$ for both data sets but because of the very long LD block, the causal location(s) could reside anywhere in this block. The rs16260 is a functional SNP within this block, which is located 365 nucleotides upstream of the transcription start site for *CDH1*. NIDDK Ileal+ data: ileal CD with involvement of at least one extra-ileal intestinal location. The blue horizontal dashed line shows the 95% Confidence Interval for the WTCCC CD dataset. The purple horizontal dashed line represents the 95% Confidence Interval for the NIDDK Combined Ileal+ data. The orange triangles represent the SNPs on the Affymetrix 500K genotyping array used in the WTCCC CD study. The blue diamonds represent the SNPs on the Affymetrix 500K genotyping array used in the WTCCC study. The orange triangles represent the SNPs on the Illumina HumanHap 300v.1 genotyping array used in the NIDDK study.

For the *IRF8* window, the WTCCC data yielded an $\hat{S}$ location 29 kb downstream of the gene, within a small block that is flanked by LD breakdown (Figure 9). The analysis of the NIDDK data for patients with any extra-ileal intestinal involvement showed a signal, which is 1.7 kb telomeric to *IRF8* within a region of LD breakdown (Figure 9).

Table 7. Association statistics and estimated location of the causal variation for two different windows covering *CDH3* (68,678-68,733 kb), *CDH1* (68,711-68,869 kb) and *IRF8* (85,933-85,956 kb) in relation to the locations on the human genome sequence.

Data	Window	Cases	χ^2_1 †	P-value	Estimated Location (kb)	95% Confidence Interval CI (kb)
WTCCC	*CDH3/CDH1*	1698	32.6	1×10^{-8}	68,746.1	68,682 – 68,836
NIDDK*	*CDH3/CDH1*	315	6.2	1×10^{-2}	68,746.1	68,684 – 68,836
WTCCC	*IRF8*	1698	34.7	4×10^{-9}	85,982.3	85,982 – 85,984
NIDDK* Jewish§	*IRF8*	38	20.9	5×10^{-6}	85,958.2	85,949 – 85,961
NIDDK* Non-Jewish§	*IRF8*	277	5.3	2×10^{-2}	85,958.2	85,935 – 85,961

*Ileal CD with involvement of at least one other extra-ileal intestinal location. §432 Jewish and 515 non-Jewish controls were used for these analyses. †A window χ^2 using composite likelihood. Note that the 95% CI for *IRF8* for the UK and North American data are non-overlapping suggesting heterogeneity of location of the causative change, i.e.: allelic heterogeneity.

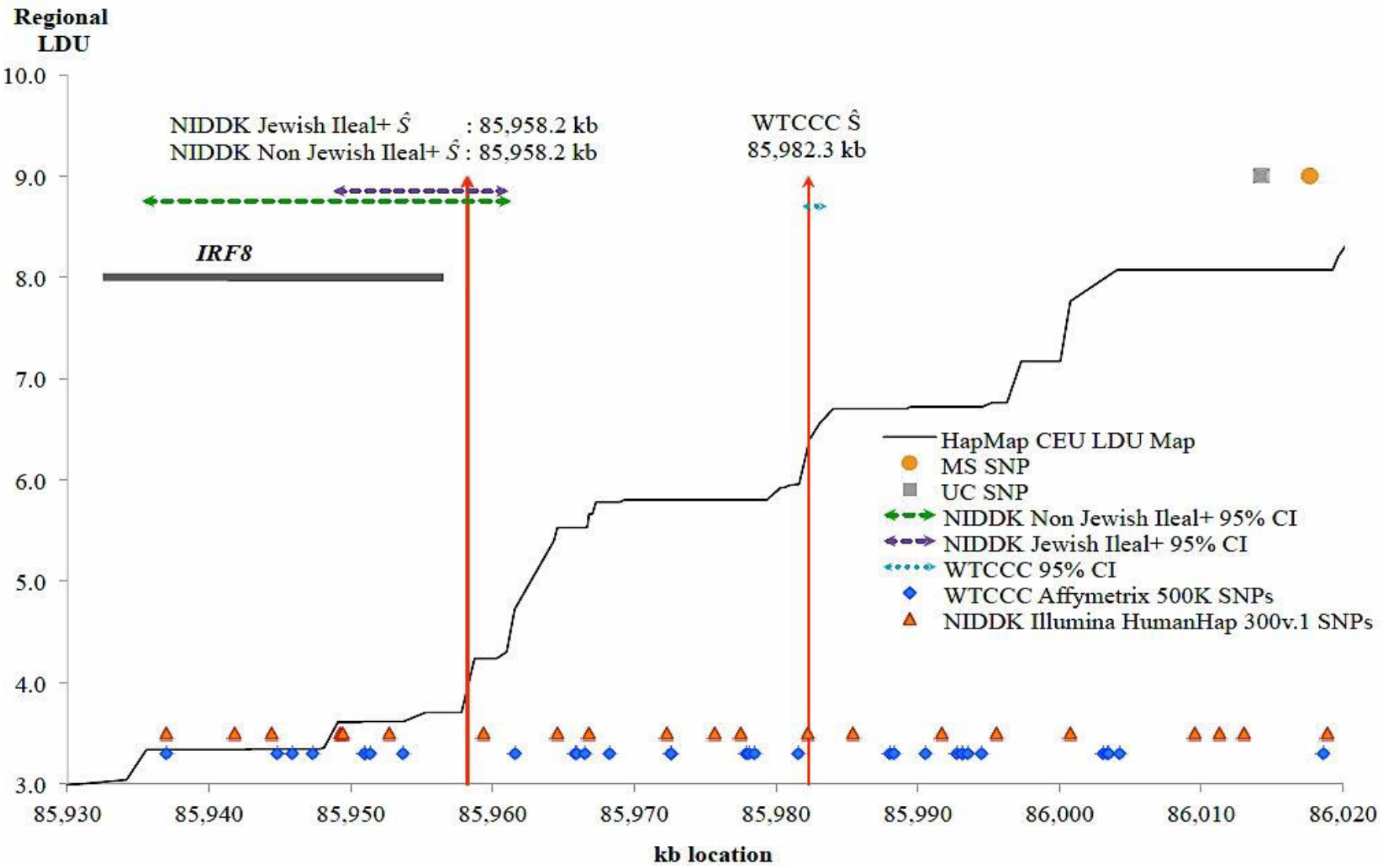

Regional
LDU
10.0
9.0
8.0
7.0
6.0
5.0
4.0
3.0
NIDDK Jewish Ileal+ Ŝ : 85,958.2 kb
NIDDK Non Jewish Ileal+ Ŝ : 85,958.2 kb
WTCCC Ŝ
85,982.3 kb
IRF8
HapMap CEU LDU Map
MS SNP
UC SNP
NIDDK Non Jewish Ileal+ 95% CI
NIDDK Jewish Ileal+ 95% CI
WTCCC 95% CI
WTCCC Affymetrix 500K SNPs
NIDDK Illumina HumanHap 300v.1 SNPs
85,930
85,940
85,950
85,960
85,970
85,980
85,990
86,000
86,010
86,020
kb location

Figure 9. Localisation within the *IRF8* region.
The LD map of the region is illustrated by the black line, which is obtained by plotting HapMap LDUs (Y axis) against kb (X axis). The red vertical arrow is the estimated location $\hat{S}$ for both data sets NIDDK Ileal+ data: ileal CD with involvement of at least one extra-ileal intestinal location. Two SNPs have been identified for UC and MS in GWA meta-analyses.The short light blue horizontal dashed line shows the 95% Confidence Interval for the WTCCC CD dataset. The purple horizontal dashed line represents the 95% Confidence Interval for the NIDDK Jewish Ileal+ data. The green horizontal dashed line represents the 95% Confidence Interval for the NIDDK Non Jewish Ileal+ data. The 95% CIs are non-symmetrical due to the block-step structure of the human genome. The blue diamonds represent the SNPs on the Affymetrix 500K genotyping array used in the WTCCC study. The blue diamonds represent the SNPs on the Affymetrix 500K genotyping array used in the WTCCC study. The orange triangles represent the SNPs on the Illumina HumanHap 300v.1 genotyping array used in the NIDDK study.

These two NIDDK datasets yielded essentially identical $\hat{S}$ locations. The 95% CI for both the Jewish and non-Jewish data included part of the *IRF8* gene with an estimated location 24 kb upstream of the WTCCC signal (Table 7, Figure 9), further suggesting allelic heterogeneity.

In this analysis 98 windows were tested. Statistical significance observed in the WTCCC data passed the Bonferroni threshold calculated based on these 98 tests (See Methods Section of this chapter). These signals were replicated using the NIDDK data with its smaller sample size and SNP coverage.

IV. DISCUSSION

Here, using the multi-marker genetic mapping approach on the available WTCCC, three novel significant signals of association with CD were found on chromosome 16q alone, all of which were replicated using the NIDDK GWAS data supporting the presence of genetic heterogeneity on chromosome 16q.

Evidence is presented for the first time that there is an independent involvement of a locus in the vicinity of *NOD2*, near the neighbouring gene *CYLD*, The pattern of association shows clear evidence of an independent signal but does not formally show that the *CYLD* gene is involved. A distant enhancer affecting *NOD2* expression could be implicated but other evidence suggests that *CYLD* itself may be important.

First, in support of this notion, genome-wide cDNA microarray analysis demonstrates that *CYLD* gene expression is down-regulated in CD (Costello et al., 2005). NOD2 interacts with Nuclear Factor- κB (NF-κB) signalling in a complex way, which includes involvement of ubiquitinylation (Abbott et al., 2004). Functional studies have so far provided conflicting evidence on the effect that the three *NOD2* mutations have with respect to inflammation and bacterial recognition (Eckmann and Karin, 2005). On the one hand, it was reported that upon bacterial LPS stimulation, the three mutant *NOD2* alleles result in the

ablation of the NOD2 Peptidoglycan-sensing activities and the failure to activate NF-κB (Bonen et al., 2003). On the other hand, a mouse-model of the frameshift 3020insC mutation showed increased levels of the cytokine IL-1β, which is a pro-inflammatory cytokine, after stimulation with MDP (Maeda et al., 2005). It is of interest that CYLD is a ***de***ubiquitinating (removal of ubiquitin) enzyme which has been shown to regulate cell proliferation, cell survival and inflammatory responses (Brummelkamp et al., 2003) and is also involved in NF-κB signalling. Dysregulation of NF-κB signalling leads to a defective immune system causing an immunodeficient or autoimmune phenotype depending on whether NF-κB function is impaired or persistent (Courtois and Gilmore, 2006).

CYLD is important in immune homoeostasis since it prevents the spontaneous activation of NF-κB in peripheral T and B lymphocytes. The peripheral T-cells from *CYLD*-deficient mice have increased sensitivity and a heightened response to T-Cell Receptor (TCR) stimulation, which leads to spontaneous inflammation in the colon (Reiley et al., 2007), and colitis-associated tumorigenesis (Zhang et al., 2006). Inflammation is the major underlying phenotype of CD and some Crohn's Disease patients develop colon cancer at a later stage in life. In this study a causal locus downstream of *CYLD* was mapped, which lies 2kb upstream of a putative gene regulatory element (Pennacchio et al., 2006) and approximately 7kb upstream of a Leprosy susceptibility GWAS SNP (Zhang et al., 2011). The element is located in an LD block spanning 8.3kb (Figure 7A, B). This regulatory element could potentially be an enhancer for *CYLD*, especially since *CYLD* expression has been found to be downregulated in CD. However, the estimated

location of the putative causal agent in this possible regulatory region could also be pointing to a mutation in which acts as a distant element in relation to *NOD2*.

The lower significance in the NIDDK data, as compared to the WTCCC result, is probably not only due to the difference in sample size but also because the phenotypic classifications were different in the two studies. The WTCCC included any subtype of CD and not just the ileal form of the condition. The 95% Confidence Interval (CI) for the *CYLD* window is very narrow for both datasets because the $\hat{S}$ is within a region of LD breakdown caused by a recombination hot spot.

Unfortunately, the data analysed in this study did not contain the actual *NOD2* mutations and the results are therefore based on data stratified according to a SNP, which is in complete LD with the three common *NOD2* mutations. Subsequent to publishing the Chromosome 16q study (Elding et al., 2011), a meta-analysis on Leprosy (Zhang et al., 2011) identified a susceptibility GWAS SNP closely located to the *CYLD* signals identified in this Chapter using the LDU multi-marker mapping approach. This is of interest since both CD and Leprosy display common immunologic characteristics, which include a Th1-cell response with granuloma formation (Zhang et al., 2009). It is also suggested that some CD cases might have a common mycobacterial cause as seen in Leprosy (Schurr and Gros, 2009). Additionally, the ENCODE (ENCyclopaedia Of DNA Elements) Consortium subsequently published a comprehensive list of all the DNA functional elements in the human genome, including Transcription Factor Binding

sites and their respective Transcription Factors as well as DNAse I Hypersensitive site, which are sites that are commonly found in transcriptionally active genes. It was interesting to find that the estimated location of the putative causal agent and the 95 % CI for the *CYLD* signal that was identified in this Chapter contained several transcription factor binding sites. Indeed the estimated location of the putative causal agent mapped precisely upstream of a binding site for the transcription factor CEBPB, in a DNase I Hypersensitive Site. The gene that codes for this transcription factor is essential in the regulation of genes that are implicated in inflammatory and immune responses (Boi et al., 2012; Mayer et al., 2013). The IBDase FP7 European consortium performed an independent conditional analysis where they stratified patients based on the actual *NOD2* mutations and subsequently identified the involvement of *CYLD* independently from that of *NOD2* in patients that do not carry the common *NOD2* mutations (I. Cleynen, 2011), which to an extent re-enforces the findings presented here. Indeed, *CYLD* is known to be modulated by some bacteria and viruses and as introduced in Chapter I of this thesis, the gut microbiota are known to play a significant role in IBD aetiology.

The second and third region were identified using the phenotypically stratified data that contained patients with ileal as well as extra-ileal involvement demonstrating that considering phenotypic heterogeneity is clearly important. The second region covers the genes *CDH3* and *CDH1* encoding cadherin proteins that participate in cell recognition, signalling, morphogenesis and tumour progression. *CDH1* encodes an epithelial cadherin (E-Cadherin) expressed in the intestine with essential functions in intestinal homeostasis (Schneider et al.,

2010). The loss of E-cadherin expression leads to apoptosis and shedding of cells and to disruption of the maturation of Paneth and Goblet cells, important in the innate immune system and microbial defence (Schneider et al., 2010). E-cadherin helps to maintain the intestinal epithelial defence system, and reduced *CDH1* expression is a feature of CD and UC patients with an inflamed intestinal epithelium (Gassler et al., 2001). A genome-wide linkage analysis on CD reported evidence of linkage on 16q, some 25cM downstream of *NOD2*, in families that did not carry the *NOD2* mutations (van Heel et al., 2003). Notably, *CDH3* and *CDH1* are at that linkage peak, approximately 25Mb downstream of *NOD2*.

However, GWA studies and meta-analyses on CD did not detect these genes despite the large number of samples and imputation based on the latest HapMap samples. A GWA study on Ulcerative Colitis on the other hand, did report *CDH1* as a new susceptibility locus (Barrett et al., 2009), with the most significant SNP for UC 180kb upstream of the gene, which will be further discussed in Chapter IV. Although *CDH1* has not been reported for CD by others to date, this analysis demonstrates that the most likely location of the causal variant is within an interval that is flanked by *CDH3* and *CDH1*. This interval is an LD block that includes a functional promoter variant (rs16260; Figure 8) (Li et al., 2000) that is also associated with IBS (Villani et al., 2010), a chronic inflammatory condition involving recurrent abdominal discomfort. *CDH1* was also detected in a GWA meta-analysis for colorectal cancer and the reported SNPs fall within our confidence interval (Houlston et al., 2008). *CDH3*, which encodes Placental-

Cadherin (P-cadherin), is also implicated in colorectal carcinomas (Milicic et al., 2008).

The third region became statistically significant after separating the data further into Jewish and non-Jewish patients and this effect shows the importance of considering heterogeneity of ancestry. The third region covers the *IRF8* gene, which encodes the transcription factor also known as Interferon Consensus Sequence-Binding Protein, which plays a negative regulatory role in cells of the immune system. Using a subset of the NIDDK data (Jewish patients with ileal CD plus additional extra-ileal involvement) yielded an estimated causal location 24 kb away from the estimate obtained using the WTCCC data. The difference in localisation may reflect allelic heterogeneity.

The significant association in the *IRF8* region on chromosome 16q was originally identified in a meta-analysis on Multiple Sclerosis (MS) (De Jager et al., 2009), which was shortly followed by a meta-analysis that identified an association between the region harbouring *IRF8* and UC (Anderson et al., 2011), the other sub-type of IBD. However, the latest association studies on CD failed to reveal any CD-specific associations on Chromosome 16q other than *NOD2*. Thus, the analysis presented here, together with other GWASs, confirm the continually emerging evidence for the presence of shared loci among different immune and inflammatory-mediated complex diseases. *IRF8* has been additionally implicated in Systemic Lupus Erythematosus (SLE) (Lessard et al., 2012) as well as MS (De Jager et al., 2009), UC (Anderson et al., 2011) and CD

(Elding et al., 2011) (Figure 9). This greatly suggests the implication of shared pathways between these diseases, and hence calls in the need for pathway analysis in complex diseases. *IRF8* expression is crucial in bone metabolism, with *IRF8*-deficient mice displaying extensive osteoporosis, one of the extra-intestinal manifestations associated with CD (Zhao et al., 2009).

This analysis provided a novel insight into the genetic and phenotypic heterogeneity of CD and demonstrates that the approach taken in this part of the project may be a promising way forward to study other frequent and complex diseases. Following the results that were obtained from this hypothesis-driven part of this project, the next step was to perform a hypothesis-free GWAS, using the LDU multi-marker mapping approach to the WTCCC and NIDDK CD datasets, which will be explored in the next chapter.

CHAPTER III. NOVEL GENE REGIONS IDENTIFIED FOR CROHN'S DISEASE.

I. INTRODUCTION

The analysis of chromosome 16q using only two GWASs described in Chapter II was very encouraging and illustrated how this LDU multi-marker approach can provide additional study power and highlighted the importance of genetic heterogeneity (*i.e.* the involvement of different risk genes in different patients) as well as that of phenotypic heterogeneity. This paved the way to studying the whole genome.

The purpose of the next part of the project was thus to analyse the whole genome using the same approach, and exploiting the positive experience of the use of phenotypic sub-classification, in order to see whether it was possible to find additional susceptibility disease genes (see the Methods section of Chapter II for further details). The work described in this chapter has been published (Elding et al., 2013).

II. METHODS

i. Genetic Analysis

As in the previous chapter, a total of 1,698 cases of CD and 2,948 ethnically matched controls were analysed from the UK WTCCC GWAS data (WTCCC, 2007) and the GWAS replication data consisted of 813 North American CD patients and 947 ethnically matched controls from the NIDDK IBDGC. Data quality control was carried out by Winston Lau as described in the previous Chapter.

Each chromosome arm was divided into non-overlapping analytical windows for both the WTCCC and the entire NIDDK GWA data-sets, as described in the Methods section of Chapter II. For each window along the whole genome, a composite-likelihood (Λ) test was used, in which all the SNPs within a window were tested simultaneously (see previous chapter for more details about the methodology used). The method returned a P-value together with an estimated location ($\hat{S}$) of the putative causal agent and the LDU estimation of $\hat{S}$ for each window is converted back to kb (NCBI build 37) by using linear interpolation of the two flanking SNPs in HapMap. In cases where $\hat{S}$ mapped to the end of a window, the window was extended in that direction to capture SNPs on the far side to improve the accuracy of location.

ii. Phenotypic Stratification

The previous analysis (See Chapter II) showed that stratifying the analysis according to the occurrence of extra-ileal intestinal involvement as opposed to ileal inflammation alone (a subdivision of which appears to be linked with genetic stratification) increased the power of detecting association of the 16q windows. Since the NIDDK GWAS data contains a detailed description of the location of patients' intestinal inflammation, the NIDDK data was stratified according to this information. For all the analytical windows that showed nominal statistical evidence of association using the WTCCC data, we analysed the same windows using the entire NIDDK data and also using a subset of the NIDDK data containing CD cases that reported extra-ileal in addition to ileal inflammation. Unlike the controls, the vast majority of cases in this subdivision were non-Jewish. We therefore analysed only non-Jewish patients with extra-ileal inflammation and non-Jewish controls. This subdivision reduced the sample size from 813 to 277 CD cases and from 947 to 515 controls. Preliminary analyses showed that fewer signals were obtained by doing this, as might be expected because of the reduced power, but our aim was to reduce as far as possible false positives obtained as a result of ethnic stratification.

iii. Previously reported and novel significant gene-regions

The first part of this analysis was to look for signals within the previously published intervals from the recent meta-analysis (Franke et al., 2010) and present P-values for the relevant data sets (See Table 8). For the novel discovery analysis a Meta-analysis P-value is presented for each analytical window that shows a) evidence of nominal statistical significance for both data samples and b) the $\hat{S}$ estimate for both samples are within 150 kb of one another. These criteria were used in order to reduce the possibility of detecting more than one signal, as a result of heterogeneity, and thus facilitate the meta-analysis of shared locations.

iv. Meta-analysis of shared locations

The significant P-values obtained from the independent datasets were combined using Fisher's combined probability test. For *k* independent tests, where each independent test of the null hypothesis results in a P-value, the P-values for each independent test can be combined as follows:

$$\chi^2 = -2\sum_{i=1}^{k} ln(P)$$

where the χ^2 has $2k$ degrees of freedom.

In order to account for multiple testing, we used a Bonferroni P-value threshold of $1x10^{-5}$, which corresponds to the total number of analytical windows and the additional analyses performed for the NIDDK data using the extra-ileal subgroup. In order to capture replications of slightly lower significance but with more precisely agreeing $\hat{S}$ values, a second threshold was used in which the $\hat{S}$ for both datasets had to lie within 80 kb of each other to be considered a replication and with a Meta P-value within the range of $1x10^{-3}$ and $1x10^{-5}$.

v. Significant gene-regions and Gene Ontology

For each of the replicated windows, the closest gene within a maximum distance of approximately 300 kb of the estimated location $\hat{S}$ for each sample was retrieved and then listed in the tables (See Tables 8 and 9). In the case where the closest gene was more than ~300kb away, the estimated location $\hat{S}$ is listed as being 'intergenic'. Using BioMart, the Gene Ontology (GO) annotations attributed to each gene were retrieved and the results are shown in Appendix II.

III. RESULTS

iv. Identifying the previously reported CD intervals

The first goal was to determine whether the LDU mapping method could detect the 71 loci previously identified from the most recent large meta-analysis of six independent datasets. Our aim was to consider the reported intervals rather than focusing on the specific genes flagged in the Consortium study. Table 8 presents the statistical evidence and location estimates, for signals within these 71 intervals (Franke et al., 2010). Since the majority of these intervals (approximately 80%) were more than 150 kb in length (defined around the most significant SNP using LD information), it was difficult to meta-analyse the results obtained using the multi-marker mapping procedure used in this analysis, since there was a risk of combining independent signals. The P-values obtained from WTCCC and NIDDK are therefore presented separately. Using the multi-marker mapping approach, a total 66 out of the 71 intervals (loci) were identified in one or both of these data sets using an uncorrected significance threshold of $P<0.05$. Only five intervals showed no evidence of association for either WTCCC or NIDDK using the LDU method. The majority of these 66 signals (88%) were identified with the larger WTCCC data with the remaining 8 intervals (12%) showing nominal evidence of association for the smaller NIDDK data set alone. Interestingly, these eight (12%) include regions (e.g. 10q22, 16p11 and 19q13)

that have previously been implicated for paediatric IBD patients (Imielinski et al., 2009). This is consistent with the fact that the NIDDK dataset probably includes a larger number of patients with early onset CD (~37%) than WTCCC, where this information is not recorded. It is worth noting that one of the 6 contributory data sets to the most recent meta-analysis also included entirely paediatric IBD patients (Franke et al., 2010). The first published results by the WTCCC (WTCCC, 2007) reported 9 loci but Table 7 shows 28 signals for WTCCC alone that passed the genome-wide significance threshold of $1x10^{-5}$.

v. Genetic Heterogeneity within the previously reported CD intervals

Unlike the recent meta-analysis, where a particular gene or genes within these intervals was identified through a series of *in silico* analyses, the genes closest to the location estimates are presented here (Table 8). In several cases these agree with the previously identified genes (e.g. *STAT3*, *IL23R*), but in other cases there are differences. For several regions two different estimates of localisation ($\hat{S}$) were obtained for WTCCC and NIDDK, both within the same interval. Heterogeneity in some regions is expected but location estimates can help identify this. For example for the interleukin-rich interval (2q12.1), the WTCCC dataset yielded an estimate within *IL18RAP* but for NIDDK the $\hat{S}$ is within *IL1RL1*. Both genes have been previously suggested as candidates within the 2q12.1 interval (Franke et al., 2010) but an independent fine mapping study has shown strong association to the *IL18RAP* rs917997 SNP for both CD and UC

(Zhernakova et al., 2008), which is approximately 10 kb away from WTCCC *Ŝ* (Figure 10). A recent study (Rivas et al., 2011) using deep sequencing identified a rare coding variant (V527L) of possible functional significance within *IL18RAP*, which is only 8 kb away from the WTCCC location, though it is probable that other functional variants are also involved. As far as the second gene is concerned, the importance of *IL1RL1* to inflammatory processes has also been documented (Akhabir and Sandford, 2010) for a variety of human pathologies including celiac disease (Amundsen et al., 2010). The LDU map clearly shows that there are several recombination hot spots between *IL18RAP* and *IL1RL1* (Figure 10) making it likely that these two signals are independent and that there is genetic heterogeneity within these previously identified intervals. This is similar to that previously noted in the case of *NOD2* and *CYLD,* in the previous chromosome 16q study (Elding et al., 2011), as described in Chapter II.

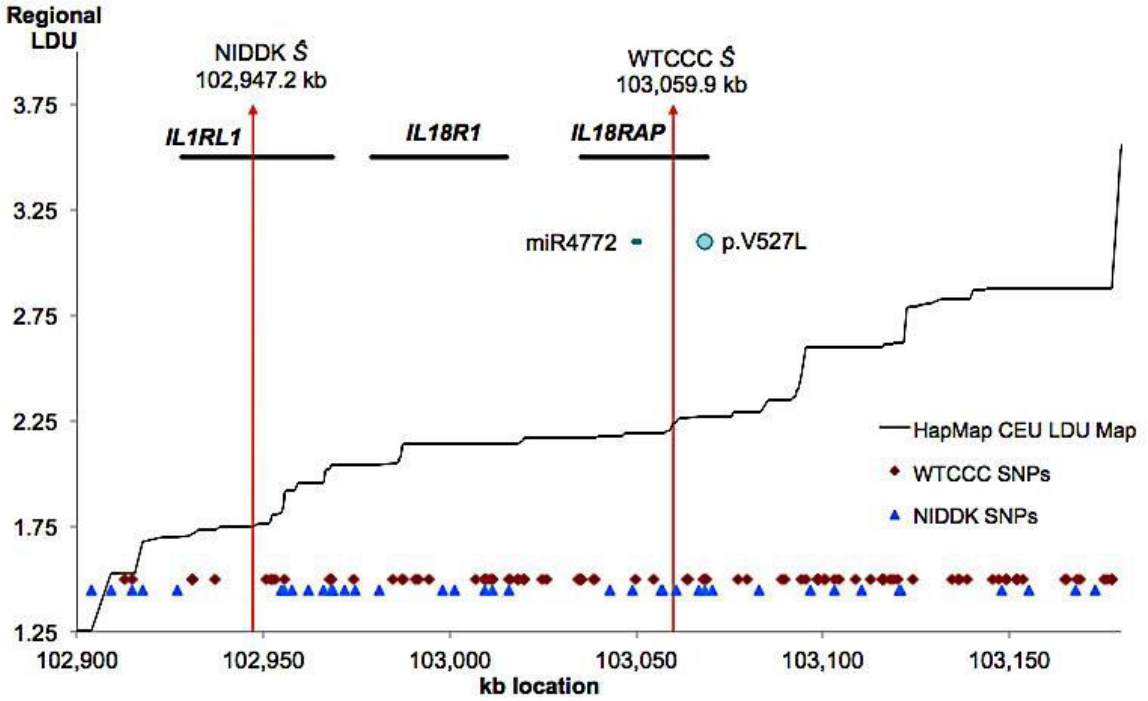


Figure 10. Localisation within the *IL1RL1/IL18RAP* region.
The LD map of the region is shown by plotting HapMap LDUs (Y axis) against kb (X axis). The red vertical arrows are the estimated locations $\hat{S}$ for the WTCCC and NIDDK datasets. The WTCCC $\hat{S}$ is within an intron and between a microRNA (miR4772) and coding rare variant for CD (p.V527L). The red diamonds represent the SNPs in this region on the WTCCC Affymetrix 500K genotyping array whereas the blue diamonds represent the SNPs in this region on the NIDDK Illumina HumanHap 300 genotyping array

vi. 134 novel gene-regions identified

In addition to the 66 localization estimates obtained, 134 novel signals were identified. Table 7 shows the list of 78 of the 134 novel signals that passed the genome-wide significance threshold of $1x10^{-5}$. Here the distance between the $\hat{S}$ locations of each signal does not exceed 150 kb, so the meta-analysis P-value is provided (as outlined in methods). Approximately 78% of the WTCCC signals were replicated using the complete NIDDK data set. However, in a third of these,

the association was significant in the complete but also in the smaller subset of the data. In this third, the smaller NIDDK extra-ileal subset showed higher significance despite the substantial decrease in the number of cases and controls. Furthermore 22% of the WTCCC signals were only replicated using this subset of the NIDDK data.

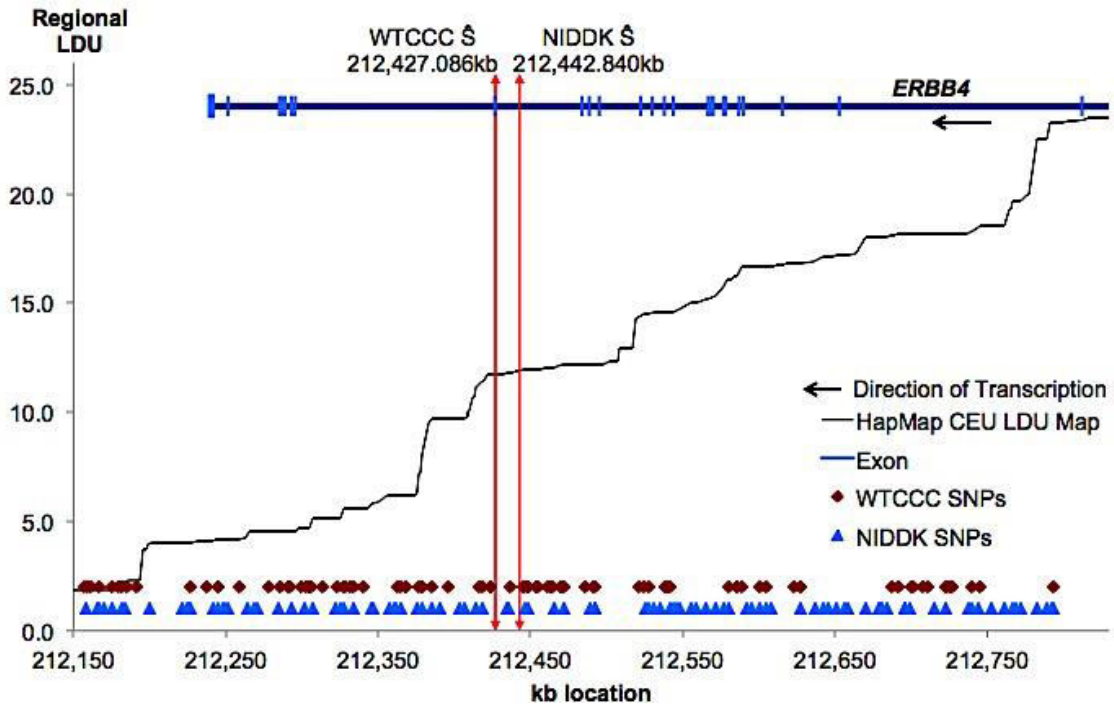


Figure 11. Localisation within the *ERBB4* region.
LD map of the region is shown by plotting HapMap Regional LDUs (Y axis) against kb (X axis). The red vertical arrows are the estimated locations $\hat{S}$ for the WTCCC and NIDDK datasets. The red diamonds represent the SNPs found in this region on the WTCCC Affymetrix 500K genotyping array whereas the blue diamonds represent the SNPs found in this region on the NIDDK Illumina HumanHap 300 genotyping array.

For the majority (74%) of the 78 signals, the location estimates from both datasets point to the same gene, with approximately half of the signals being intragenic and half intergenic locations (gene with asterisk implies that at least one of the signals is within the gene (Table 9 and 10). In all cases, including the ~50% that reside outside of genes, the closest gene is considered as a candidate. It should however be noted that in some cases long range cis-elements may be involved, which regulate genes far outside the 150 kb distance, with other genes intervening, or regulatory elements may exist within an intron of an adjacent gene (Jones and Swallow, 2011). Indeed, cis-acting gene expression Quantitative Trait Loci (eQTL) have been implicated in both Mendelian and complex disorders. eQTLs are loci in the genome that control mRNA expression levels. There have been several efforts in correlating GWAS disease-risk loci with eQTL genes in an attempt to further explain the missing heritability that is very common to complex conditions, including CD. In fact, some GWAS CD susceptibility SNPs have been shown to be associatd with variable expression levels of *PTGER4, CARD9*, *ERAP2*, *TNFSF11* in Epstein-Barr Virus (EBV) transformed Lympholastoid Cell Lines (LCLs) (Libioulle et al., 2007; Montgomery and Dermitzakis, 2011).

Several of the novel genes have been reported to be involved in inflammatory/ immune dysregulation conditions (*e.g. DOCK8, ITGA9*). In other cases the genes have been previously implicated in colonic inflammation or colonic cell morphology based on functional studies. By way of illustration, Figure 11 presents an example of the estimated locations ($\hat{S}$) for the genomic region that harbours *ERBB4* (2q34 window). For both datasets, this window was significantly

associated with CD with a Meta P-value of 3×10^{-6} (Table 9). Figure 11 shows the LDU map starting from the 3' region and spanning half way down the gene, which stretches across more than 1 Megabase. The map shows numerous short LD blocks across the region. However, despite the large LD breakdown in the area, both estimates of $\hat{S}$ are close to each other and within the same intron.

Table 10 shows the list of 56 of the 134 novel signals that showed association with CD where the P-values ($1x10^{-3}$ – $1x10^{-5}$range) based upon meta-analyses were below the Bonferroni threshold but estimates of localization in the two data sets were not more than 80 kb apart from each other. Most of the signals (96%) in Table 10, give estimates of $\hat{S}$ for the same gene.

IV. DISCUSSION

Using the LD approach 66 of the 71 previously reported loci from meta-analysis were confirmed and 134 novel gene regions that are associated with Crohn's Disease have been identified, providing evidence for 200 gene regions that include CD susceptibility loci. More precise location estimates are provided for the 66 previously published intervals, as well as locations for the 134 new signals. The very large numbers of genes listed in Table 9 and 10 support the idea that CD is truly polygenic and complex in nature and that much of this complexity is genetic in origin. Many genes show functions that are compatible with involvement in immune/inflammatory processes as well as integrity of the intestinal epithelium and differentiation.

In the CD analysis described in this chapter, many of the replicated signals were identified only, or more strongly, using a subset of the NIDDK data, which contained patients with ileal CD who had involvement of at least one extra-ileal intestinal location. Despite the substantial decrease in sample size, analysis of this sub-phenotype using this method yielded much higher power than the analysis of the full data. This indicates that phenotypic heterogeneity is clearly important and that accurate as well as detailed phenotype information is crucial in genetic studies, especially when assessing a trait showing clinical variation.

Rather than mapping to LD intervals around the most significant SNP, this method provides estimated locations of the putative causal agents. These location estimates are of great importance since they do not only pinpoint precise locations but also provide the possibility of detecting possible heterogeneity, as shown with the example of *NOD2/CYLD* on chromosome 16q described in the previous Chapter and with the example of *IL1RL1/IL18RAP* on chromosome 2q identified here .

Several of the genes that were identified in this part of the analysis had been previously implicated in IBD through functional studies. An example is *ERBB4* (which encodes Receptor protein-tyrosine kinase erbB-4) on 2q34 (See Table 9). *ERBB4* is expressed at high levels in the inflamed colonic mucosa of CD patients (Frey et al., 2009). Using adult mouse colon, it was also shown that ERBB4 is an important regulator in the epithelial response to inflammation and injury (Frey et al., 2009). *ERBB4* expression has been linked to a number of cellular processes such as cell survival, proliferation and tumorigenesis in different tissues (Erlich et al., 2001; Starr et al., 2006). A recent study suggested that this elevated *ERBB4* expression could lead to colitis-associated development of colorectal tumours (Frey et al., 2010). *ERRB4* has also been implicated, together with E-cadherin (encoded by *CDH1* on 16q22.1, previously identified as a risk gene for CD and ulcerative colitis), in the suppression of anoikis (Kang et al., 2007) (programmed cell death as a result of cell detachment from the extracellular matrix), suggesting that there may be interaction between these two genes. Despite the fact that *ERBB4* is a very large gene (>1 Mb), the estimated locations of the putative causal agent for both datasets (WTCCC and NIDDK) are very close and within

the same intron, making it an excellent target for follow-up fine mapping studies and re-sequencing.

Several of the genes listed in Table 10, which presents 56 gene regions with lower confidence of association, but with very close location estimates between the two datasets, are also very interesting. For example, the window harbouring the *BTLA* gene gave location estimates that are just 2 kb apart. A study using a mouse model of colitis has shown functional evidence of *BTLA* involvement in colitis (Steinberg et al., 2008). Box 1 illustrates a functional overview of some genes of interest that were identified in this part of the project. The molecular and functional attributes of all the genes identified in this chapter are illustrated in Appendix II.

The LDU multi-marker mapping approach successfully identified many additional novel CD susceptibility loci by re-analysing publically available GWAS data. The next step would be to shift the focus towards UC and identify whether, using the LDU multi-marker mapping approach, more loci could be detected for this-subtype, which is known to have a lower genetic contribution than UC.

Box 1: Functional overview of some genes of interest identified.

***ERBB4*:** (v-erb-a erythroblastic leukemia viral oncogene homologue 4). This gene is a member of the Tyrosine protein kinase family and the Epidermal Growth Factor receptor subfamily. ErbB4 is a TNF-inducible receptor and has been found to regulate repair in response to injury and inflammation. Activation of E-Cadherin (encoded by *CDH1*- See Table 3) which suppresses anoikis (a type of apoptosis involving detachment from the extracellular matrix) activates ErbB4 receptor tyrosine kinase and subsequent activation of PI3 Kinase pathway leading to cell proliferation.

***BTLA*:** (B and T Lymphocyte Associated). BTLA is a member of the immunoglobulin superfamily containing a single Ig domain. It is also a receptor that relays inhibitory signals to suppress the immune response. BTLA engagement is required to prevent colitis acceleration in a mouse model. The expression of *TNFSF14* and *BTLA* on both adaptive and innate cells strongly indicates that this ligand-receptor system could play several roles in immunity and inflammation. Mutations in this gene have been associated with an increased risk of Rheumatoid Arthritis.

***ITGA9*:** (integrin alpha 9). This gene encodes an alpha integrin. Integrin are heterodimeric integral membrane glycoproteins composed of an alpha chain and beta chain that mediate cell-cell and cell-matrix adhesions. ITGA9 is normally active in colonic epithelia. Alpha-9 deficient bone marrow cells in mice show a marked decrease in STAT3 phosphorylation (See Table 1 on chromosome 17q) after G-CSF stimulation in granulopoeisis.

***DOCK8*:** (dictator of cytokinesis 8). This gene encodes a member of the DOCK180 family of Guanine Nucleotide Exchange Factors. Mutations in this gene result in the autosomal recessive hyperimmunoglobulin E (hyper-IgE) syndrome characterized by immunodeficiency, lack of connective tissue and recurrent infections among other symptoms. *DOCK8* is most highly expressed in B and T lymphocytes and mutations in *DOCK8* diminish clonal expansion of both CD4 and CD8 T cells after engagement of TCR and co-stimulatory molecules. It was found that patients with large homozygous deletions in *DOCK8* had IBD or suggestive IBD. *DOCK8* is also required to form a normal T-cell immune-synapse with antigen-presenting Dendritic Cells (DCs).

***BRD7*:** (bromodomain containing 7). The product of this gene has been identified as a component of one form of SWI/SNF chromatin remodeling complex which also interacts with p53. BRD7 depletion delays replicative senescence and extends cellular life-span.

***TEC*:** (Tec protein tyrosine kinase). Tec is part of the Tec family non-receptor protein tyrosine kinases. Tec kinase is an central component of T-cell signaling and has a distinct role in T cell activation. This gene plays a role in Calcium influx, apoptosis, gene expression, actin re-organisation as well as cell-adhesion and migration and is found to be activated by bacterial Lipopolysaccharide (LPS).

***USP1*:** (ubiquitin-specific peptidase 1). This gene encodes a member of the ubiquiting specific processing (UBP) family of proteases that is a deubiquitinating (DUB) enzyme. PCNA (a central component of the ubiquitin-based molecular switch dictating error-free versus error-prone DNA repair) polyubiquitination is negatively regulated by USP1 in the absence of DNA damage.

Table 8. Whole-genome association statistics and closest gene to the estimated location $\hat{S}$ of the causal variant for the previously reported 71 LD intervals.

Chr	Reported Interval (Mb)	WTCCC $\hat{S}$	WTCCC P-value	NIDDK $\hat{S}$	NIDDK P-value	Gene (WTCCC/ NIDDK If different)
1p31.3	67.36 - 67.77	67,684.8	2.1×10^{-11}	67,707.3	3.0×10^{-10}	*IL23R**
1p13.2	113.95 - 114.62	114,560.3	1.4×10^{-03}	114,108.3	5.0×10^{-03}	*OLFML3/MAGI3**
1p36.23	7.74 - 7.97	-	-	7,801.1	2.2×10^{-05}	*CAMTA1**
1q23.3	160.69 - 162.47	160,887.0	2.6×10^{-03}	160,837.3	5.8×10^{-03}	*ITLN2/CD244*
1q25.1	172.66 - 172.95	172,891.9	3.6×10^{-13}	-	-	*TNFSF18*
1q32.1	200.85 - 201.06	200,877.3	3.0×10^{-14}	200,957.2	1.5×10^{-02}	*C1orf106*/KIF21B**
1q22	154.97 - 156.13	-	-	-	-	-
1q31.3	197.32 - 197.95	197,786.1	1.2×10^{-04}	-	-	*DENND1B*
1q32.1	206.80 - 207.03	-	-	-	-	-
2p23.3	25.45 - 25.60	25,389.2	3.5×10^{-02}	-	-	*POMC**
2p23.3	27.39 - 27.86	27,697.4	5.4×10^{-04}	27,619.3	2.3×10^{-11}	***IFT172*/PPM1G****
2p21	43.45 - 43.95	43,840.8	2.0×10^{-05}	43,581.0	1.1×10^{-02}	*THADA**
2p15	60.92 - 61.89	61,395.5	1.9×10^{-05}	-	-	*C2orf74*
2q37.1	234.15 - 234.57	234,144.8	5.6×10^{-25}	234,171.9	2.2×10^{-09}	*ATG16L1**
2q12.1	102.80 - 103.30	103,059.8	1.8×10^{-08}	102,951.7	1.7×10^{-05}	*IL18RAP*/IL1RL1**
2q33.1	198.14 - 198.96	198,957.7	1.6×10^{-02}	198,205.8	1.1×10^{-03}	*PLCL1*/ANKRD44*
2q37.1	231.05 - 231.23	231,109.3	2.8×10^{-05}	-	-	*SP140**
3p21.31	48.18 - 51.75	49,743.9	1.7×10^{-31}	48,818.0	3.3×10^{-19}	*RNF123*/PRKAR2A**
3p24.3	18.60 - 18.88	18,707.0	5.3×10^{-03}	18,629.1	2.6×10^{-05}	*SATB1*
5p13.1	39.84 - 40.96	40,447.1	6.0×10^{-51}	40,288.0	2.7×10^{-07}	*PTGER4*
5q31.1	129.38 - 132.02	131,748.2	2.7×10^{-22}	131,631.4	8.9×10^{-07}	*C5orf56*/SLC22A4**
5q33.1	150.03 - 150.40	150,230.7	3.6×10^{-09}	-	-	*IRGM*
5q33.3	158.50 - 158.95	158,825.7	5.6×10^{-06}	158,826.0	3.5×10^{-02}	*IL12B*
5q13.2	72.45 - 72.58	72,465.4	5.8×10^{-03}	-	-	*TMEM174*
5q15	96.08 - 96.42	96,372.2	1.3×10^{-08}	-	-	*LNPEP*
5q31.3	141.41 - 141.64	141,480.3	1.3×10^{-02}	-	-	*NDFIP1*
5q35.2	173.22 - 173.54	-	-	173,417.9	6.7×10^{-03}	*C5orf47**
6p22.3	20.49 - 21.14	20,559.3	4.3×10^{-03}	20,433.1	2.8×10^{-03}	*CDKAL1*/E2F3**
6p21.32	31.38 - 32.87	32,720.6	1.2×10^{-27}	32,193.4	1.6×10^{-07}	*HLA-DQB2*
6p25.2	3.41 - 3.47	3,417.6	1.5×10^{-07}	-	-	*SLC22A23**
6q21	106.39 - 106.56	-	-	106,421.7	4.0×10^{-02}	*PRDM1*
6q27	167.34 - 167.55	167,441.4	3.3×10^{-09}	167,355.9	8.5×10^{-03}	*FGFR1OP*/RNASET2**
6q15	90.80 - 91.08	-	-	-	-	-
6q25.3	159.34 - 159.54	159,539.7	1.2×10^{-03}	-	-	*FNDC1*
7p12.2	50.25 - 50.33	50,255.3	9.4×10^{-04}	-	-	*C7orf72*
8q24.13	126.47 - 126.58	126,544.1	4.5×10^{-05}	-	-	*TRIB1*
8q24.21	129.49 - 129.60	129,567.2	1.3×10^{-04}	129,570.7	2.9×10^{-03}	***intergenic***
9p24.1	4.94 - 5.30	5,274.6	1.4×10^{-04}	5,270.3	5.4×10^{-03}	***RLN2***
9q32	117.43 - 117.70	117,555.6	3.9×10^{-04}	-	-	*TNFSF15*
9q34.3	139.13 - 139.42	139,280.8	2.3×10^{-05}	139,299.7	2.6×10^{-03}	*SNAPC4*/SDCCAG3**
10p11.21	35.18 - 35.90	35,554.8	3.4×10^{-29}	-	-	*CCNY**
10p15.1	6.03 - 6.17	6,189.9	3.6×10^{-03}	6,165.5	1.2×10^{-08}	*PFKFB3*/RBM17**
10q21.2	64.30 - 64.76	64,448.3	1.2×10^{-18}	-	-	*ZNF365*
10q24.2	101.27 - 101.34	101,324.7	1.2×10^{-32}	101,274.6	2.3×10^{-02}	*NKX2-3*
10q21.1	59.83 - 60.14	59,892.8	9.3×10^{-09}	59,893.1	1.6×10^{-02}	*IPMK*
10q22.3	81.00 - 81.10	-	-	81,019.8	1.5×10^{-02}	*ZMIZ1**
11q13.5	76.02 - 76.36	76,304.4	4.4×10^{-10}	-	-	*C11orf30*
11q12.2	61.52 - 61.68	61,680.2	3.4×10^{-02}	-	-	*RAB3IL1**
11q13.1	63.82 - 64.29	64,027.0	5.8×10^{-03}	64,137.4	3.2×10^{-02}	*PLCB3*/RPS6KA4**
12q12	40.13 - 41.02	40,598.3	3.0×10^{-02}	40,368.0	1.7×10^{-03}	*LRRK2/SLC2A13**
13q14.11	44.23 - 44.64	44,491.1	6.6×10^{-12}	44,607.7	3.0×10^{-02}	*LACC1/LINC00284*
13q14.11	42.82 - 43.10	43,022.6	1.3×10^{-07}	-	-	*TNFSF11*

Table 8 Cont.

Chr	Reported Interval (Mb)	WTCCC $\hat{S}$	WTCCC P-value	NIDDK $\hat{S}$	NIDDK P-value	Gene (WTCCC/ NIDDK If different)
14q24.1	69.16 - 69.32	69,199.5	4.1×10^{-02}	-	-	*ZFP36L1*
14q31.3	88.21 - 88.64	88,364.3	6.2×10^{-05}	88,216.2	5.1×10^{-03}	*GALC*
15q22.33	67.41 - 67.48	67,410.8	1.5×10^{-02}	-	-	*SMAD3**
16p11.2	28.29 - 29.03	-	-	28,940.3	4.5×10^{-02}	*CD19*
16q12.1	50.46 - 50.85	50,803.2	2.6×10^{-15}	50,803.2	1.1×10^{-09}	*NOD2**
		50,846.3	1.6×10^{-13}	50,846.5	4.1×10^{-04}	*CYLD*
17q12	37.37 - 38.26	37,985.9	1.1×10^{-03}	37,888.9	4.2×10^{-05}	*IKZF3*/MIEN1*
17q21.2	40.32 - 41.00	40,501.5	2.0×10^{-18}	-	-	*STAT3**
17q12	32.49 - 32.68	-	-	32,634.2	4.1×10^{-02}	***CCL8***
18p11.21	12.74 - 12.93	12,875.2	2.9×10^{-12}	-	-	*PTPN2**
19p13.3	1.09 - 1.18	1,127.7	2.0×10^{-02}	-	-	*SBNO2**
19p13.2	10.40 - 10.64	10,560.9	1.3×10^{-02}	-	-	*PDE4A**
19q13.11	33.73 - 33.78	-	-	-	-	-
19q13.33	49.09 - 49.28	-	-	49,212.0	3.6×10^{-04}	*FUT2*
20q13.33	62.18 - 62.48	62,327.9	2.3×10^{-04}	-	-	*RTEL1-TNFRSF6B**
21q12.1	16.70 - 16.85	16,841.4	4.4×10^{-07}	-	-	*USP25*
21q22.3	45.59 - 45.70	45,609.1	5.5×10^{-06}	-	-	*ICOSLG*
21q11.21	21.81 - 22.06	-	-	-	-	-
22q12.2	29.90 - 30.67	30,228.8	8.0×10^{-07}	-	-	*ASCC2**
22q13.1	39.67 - 39.81	-	-	39,686.8	1.1×10^{-02}	*RPL3*

Bold denotes that the signal is replicated using the data taken only from the Non-Jewish 'Ileal' patients who also had extra-ileal inflammation; *at least one of the estimated locations $\hat{S}$ is located within the identified gene.

Table 9. Whole-genome assocation statistics and the closest gene to the estimated location $\hat{S}$ of the causal variant for the convincingly confirmed 78 gene-regions.

Chr	WTCCC $\hat{S}$	NIDDK $\hat{S}$	Gene (WTCCC/NIDDK if different)	Meta-P-value
1p31.3	61,982.8	61,938.8	*NFIA*[§]	1.4×10^{-07}
1p31.3	64,095.1	64,005.9	***PGM1*/EFCAB7***[§]	1.1×10^{-08}
1q21.3	151,788.8	151,837.4	*RORC*/THEM4*	1.2×10^{-09}
1q42.2	233,765.3	233,835.4	***KCNK1***[§]	5.3×10^{-08}
1q43	236,709.1	236,684.6	***LGALS8****	2.8×10^{-07}
1q43	236,862.0	236,805.4	***ACTN2*/HEATR1***	1.3×10^{-10}
2p25.2	6,394.6	6,425.5	*LOC400940*	5.1×10^{-09}
2p24.2	16,442.2	16,452.4	***FAM49A***	5.6×10^{-06}
2p23.3	25,247.9	25,138.3	*DNAJC27-AS1*/ADCY3**	8.7×10^{-18}
2p13.1	74,616.6	74,659.7	*DCTN1*/RTKN**[§]	4.9×10^{-22}
2q14.1	117,951.4	117,821.6	***intergenic***	3.3×10^{-06}
2q24.3	168,944.1	168,810.8	*STK39**	3.3×10^{-07}
2q32.1	185,889.2	185,791.4	*ZNF804A**	7.4×10^{-09}
2q34	212,427.1	212,442.8	*ERBB4**	3.1×10^{-06}
3p22.3	36,213.9	36,097.7	*STAC/ARPP21*	2.6×10^{-09}
3p22.2	37,598.6	37,602.5	***ITGA9****	6.2×10^{-09}
3q11.2	96,376.9	96,304.5	***EPHA6***[§]	6.6×10^{-09}
3q12.1	99,338.2	99,330.9	*COL8A1*	1.2×10^{-07}
4p12-11	48,197.8	48,342.9	*TEC*/SLAIN2*	3.4×10^{-15}
4q12	54,551.5	54,668.6	*LNX1/RPL21P44*[§]	2.7×10^{-06}
4q12	55,260.5	55,241.3	***PDGFRA***[§]	2.9×10^{-09}
4q25	110,193.1	110,128.2	***COL25A1***[§]	6.8×10^{-06}
4q35.1	186,241.5	186,267.1	*SNX25**	5.2×10^{-08}
5q22.2	112,590.6	112,572.4	***MCC***[§]	8.5×10^{-06}
5q23.2	126,822.9	126,898.0	*MEGF10/PRRC1*[§]	5.5×10^{-10}
6p22.3	24,443.8	24,457.3	***GPLD1****	4.0×10^{-07}
6p22.1	29,605.2	29,575.4	*GABBR1**	5.5×10^{-06}
6q12	65,246.8	65,244.5	*EYS**	1.6×10^{-07}
6q25.3	159,670.8	159,794.4	*FNDC1**	1.5×10^{-08}
6q27	166,505.9	166,501.1	*T*[§]	7.0×10^{-07}
7p21.1	20,044.2	19,975.1	***MACC1/TMEM196***	1.9×10^{-07}
7p15.3	24,175.9	24,271.9	*NPY*[§]	3.9×10^{-07}
7p14.1	41,898.3	41,939.6	***INHBA-AS1/GLI3***	2.1×10^{-06}
7q21.11	80,401.5	80,470.9	*SEMA3C**	4.8×10^{-06}
7q31.31	120,409.2	120,337.3	*TSPAN12/KCND2**	8.7×10^{-06}
7q31.33	125,266.9	125,219.7	***Intergenic***[§]	3.7×10^{-10}
7q32.3	131,467.2	131,394.6	*PODXL*[§]	7.8×10^{-07}

Table 9 continued

Chr	WTCCC $\hat{S}$	NIDDK $\hat{S}$	Gene (WTCCC/NIDDK if different)	Meta-P-value
8p22	17,632.9	17,635.3	*MTUS1**	$5.5x10^{-07}$
8q12.1	57,504.1	57,507.6	*LOC100507632*	$5.2x10^{-06}$
8q21.13	80,385.0	80,277.0	*STMN2*[§]	$1.3x10^{-12}$
8q21.13	83,670.2	83,786.5	***intergenic***	$1.9x10^{-08}$
8q22.3	104,465.6	104,363.8	*DCAF13/FZD6*[§]	$4.4x10^{-06}$
8q23.3	114,908.3	114,815.4	***intergenic***[§]	$4.9x10^{-11}$
9p24.3	396.8	468.1	*DOCK8**[§]	$2.6x10^{-07}$
10p12.1	25,571.8	25,686.6	***GPR158****[§]	$1.5x10^{-09}$
10q21.1	61,019.5	61,135.2	***FAM13C****	$2.2x10^{-06}$
10q21.2	63,604.2	63,553.3	*ARID5B/C10orf107*	$7.3x10^{-09}$
10q22.2	75,697.5	75,698.5	***PLAU***[§]	$2.6x10^{-13}$
10q22.3	79,392.4	79,247.9	*KCNMA1**	$2.5x10^{-06}$
10q25.1	111,613.9	111,622.3	*XPNPEP1*	$1.8x10^{-06}$
10q25.3	115,370.8	115,381.2	***NRAP****	$1.0x10^{-05}$
11p15.4	5,785.5	5,836.4	*OR52N4/OR52N2*[§]	$3.3x10^{-10}$
11p14.1	27,827.8	27,790.8	*BDNF*	$8.2x10^{-07}$
11q12.1	56,626.0	56,625.0	***OR9G4***	$1.0x10^{-05}$
11q24.2	126,774.5	126,749.9	***KIRREL3****	$1.1x10^{-06}$
12p11.22	28,744.2	28,741.7	*CCDC91*[§]	$1.1x10^{-06}$
12q12	44,676.8	44,823.7	***TMEM117****[§]	$2.4x10^{-06}$
12q22	92,998.8	93,090.8	***C12orf74***[§]	$7.6x10^{-09}$
12q23.1	101,350.2	101,350.1	***ANO4****	$9.3x10^{-06}$
13q14.3	52,081.3	52,078.0	***MIR4703***	$8.9x10^{-08}$
14q21.1	39,671.6	39,601.9	*PNN/GEMIN2**[§]	$6.4x10^{-07}$
14q31.1	80,015.5	80,139.6	***NRXN3****[§]	$7.6x10^{-07}$
15q13.3	31,839.1	31,737.8	*OTUD7A**	$1.0x10^{-08}$
15q14	34,414.0	34,291.8	*PGBD4/CHRM5**	$8.9x10^{-06}$
16p13.3	7,262.9	7,200.6	*RBFOX1*[§]	$1.0x10^{-05}$
16p13.3	7,272.8	7,391.8	***RBFOX1****	$3.6x10^{-06}$
16q12.1	50,417.0	50,417.2	*BRD7*	$2.0x10^{-08}$
16q22.1	67,464.1	67,422.7	***HSD11B2***[§]	$7.5x10^{-08}$
16q24.1	85,982.3	85,958.0	***IRF8***	$1.5x10^{-08}$
17q25.1	73,344.9	73,411.5	*GRB2*/MIR3678*	$4.3x10^{-14}$
18q11.2	19,677.2	19,620.4	*GATA6*	$2.6x10^{-07}$
18q12.3	40,517.4	40,666.0	***RIT2****[§]	$6.0x10^{-18}$
18q21.1	44,940.9	44,800.7	*IER3IP1*[§]	$7.8x10^{-06}$
20p12.3	7,901.0	7,878.5	*HAO1**[§]	$1.4x10^{-07}$
20q12	37,960.3	38,106.0	*LOC339568*	$3.9x10^{-06}$
20q13.12	44,741.5	44,672.0	***CD40/SLC12A5****[§]	$7.7x10^{-08}$
22q12.3	37,325.9	37,349.3	*CSF2RB*/LOC1005006241*	$2.6x10^{-08}$
Xq28	147,228.4	147,114.7	*FMR1NB*	$4.4x10^{-07}$

Bold denotes that the signal is replicated using the data taken only from the Non-Jewish 'Ileal' patients who also had extra-ileal inflammation; *at least one of the estimated locations $\hat{S}$ is located within the identified gene;[§]the signal is significant in the Pooled NIDDK data as well as the stratified Non Jewish Ileal and extra-ileal dataset.

Table 10. Whole-genome association statistics and the closest gene to the estimated location $\hat{S}$ of the causal variant for the confirmed 56 gene-regions (P-value 10^{-3}-10^{-5}).

Chr	WTCCC $\hat{S}$	NIDDK $\hat{S}$	Gene (WTCCC/NIDDK if different)	Meta-P-value
1p34.1	45,764.3	45,764.4	*LOC400752*	$3.6x10^{-04}$
1p32.2	58,535.6	58,575.4	*DAB1**	$2.7x10^{-04}$
1p31.3	68,056.4	68,072.2	***GADD45A***	$2.9x10^{-05}$
1p31.1	83,125.7	83054.7	***intergenic***	$1.4x10^{-04}$
1q31.3	194,273.4	194,332.3	*intergenic*[§]	$2.2x10^{-04}$
1q32.1	204,129.7	204,143.0	*REN**	$3.4x10^{-04}$
1q42.12	225,508.2	225,554.8	*DNAH14**	$3.0x10^{-05}$
1q42.12	226,248.8	226,249.4	***H3F3A***[§]	$9.1x10^{-04}$
2p14	67,869.6	67,871.1	*ETAA1*[§]	$2.1x10^{-05}$
2q11.2	102,625.1	102,625.1	*IL1R2**	$4.9x10^{-04}$
2q31.1	170,038.8	170,037.2	***LRP2****	$8.4x10^{-04}$
2q36.3	226,258.2	226,258.4	*NYAP2*	$1.2x10^{-05}$
3p14.3	55,192.0	55,115.1	*CACNA2D3*	$5.7x10^{-05}$
3p12.3	77,377.2	77,377.3	*ROBO2**	$8.2x10^{-04}$
3q13.2	112,181.4	112,184.1	*BTLA**	$2.9x10^{-05}$
3q24	148,459.5	148,419.7	*AGTR1**	$2.1x10^{-04}$
4p15.33	13,075.6	13,050.3	*HSP90AB2*[§]	$1.5x10^{-04}$
4q13.2	67,799.9	67,792.9	***intergenic***	$8.2x10^{-04}$
4q27	121,797.6	121872.1	*PRDM5**	$2.1x10^{-05}$
5p14.3	23,042.2	23,002.6	***CDH12***	$3.9x10^{-05}$
5q21.1	100,970.9	100,970.8	***intergenic***	$6.0x10^{-04}$
5q33.3	156,594.7	156,627.7	*FAM71B/ITK**[§]	$8.2x10^{-05}$
6p22.3	21,462.0	21,467.8	*SOX4*	$5.6x10^{-04}$
6q21	107,433.4	107,473.5	*BEND3**	$3.6x10^{-05}$
6q21	113,971.2	113,977.4	*MARCKS*	$7.1x10^{-05}$
6q22.31	125,214.4	125,215.3	*STL*[§]	$3.0x10^{-04}$
6q23.3	137,546.9	137,565.2	*IFNGR1*	$3.5x10^{-04}$

Table 10 continued

Chr	WTCCC $\hat{S}$	NIDDK $\hat{S}$	Gene (WTCCC/NIDDK if different)	Meta-P-value
6q25.3	156,152.9	156,190.7	*intergenic*	1.9×10^{-04}
6q25.3	158,344.5	158,365.2	*SNX9*§*	3.1×10^{-04}
7p21.1	20,286.7	20,301.2	***MACC1***	5.5×10^{-05}
7p14.3	34,739.3	34,757.7	***NPSR1****	3.2×10^{-04}
7q31.32	122,995.2	122,993.0	***IQUB***§	1.6×10^{-04}
7q34	142,653.2	142,650.9	*KEL**	1.1×10^{-04}
8q24.22	134,050.4	134,063.7	*SLA**	1.5×10^{-04}
9q21.13	74,603.0	74678.4	*C9orf85/C9orf57*	7.7×10^{-4}
9q21.2	80,206.6	80,206.7	***GNA14****§	2.7×10^{-04}
9q33.1	121,364.8	121,365.0	***intergenic***	5.1×10^{-04}
10q23.31	90,787.7	90,787.7	*FAS*§	2.7×10^{-04}
11p15.5	1,539.0	1,539.0	*MOB2**	2.1×10^{-04}
11p15.3	12,387.3	12,331.6	***MICALCL****	5.1×10^{-05}
12q14.3	67,645.2	67,640.5	***CAND1***	6.9×10^{-04}
12q21.32	89,154.8	89,148.0	*KITLG*§	5.7×10^{-05}
12q24.32	126,746.8	126,753.9	*LOC100128554*§	9.9×10^{-04}
13q12.13	27,389.5	27,391.1	***GPR12***	1.3×10^{-04}
13q13.3	36,090.9	36,159.4	*NBEA**	2.7×10^{-04}
13q31.3	90,018.8	90,018.6	***intergenic***	7.5×10^{-04}
13q31.3	90,206.6	90,176.3	***intergenic***	6.9×10^{-04}
14q31.1	82,444.3	82,452.3	*intergenic*§	5.4×10^{-05}
15q21.2	50,193.4	50,171.3	***ATP8B4****	8.8×10^{-04}
16q21	61,777.3	61,777.1	*CDH8**	1.4×10^{-05}
16q23.1	76,295.6	76,275.6	*CNTNAP4*	9.1×10^{-04}
16q23.1	78,458.5	78,410.4	*WWOX**§	6.6×10^{-04}
20p21.1	15,802.7	15,876.4	***MACROD2****	7.1×10^{-04}
21q22.2	41,532.5	41,513.2	*DSCAM**	7.3×10^{-04}
22q12.3	35,540.3	35,540.4	*ISX*§	2.8×10^{-04}
Xq23	115,981.3	115,966.6	***intergenic***	7.3×10^{-05}

Bold denotes that the signal is replicated using the data taken only from the Non-Jewish 'Ileal' patients who also had extra-ileal inflammation; *at least one of the estimated locations $\hat{S}$ is located within the identified gene; §the signal is significant in the Pooled NIDDK data as well as the stratified Non Jewish Ileal and extra-ileal dataset.

CHAPTER IV. NOVEL GENE REGIONS IDENTIFIED FOR ULCERATIVE COLITIS.

I. INTRODUCTION

Following the success in identifying novel genes for Crohn's disease using the LDU mapping approach, the attention was shifted towards Ulcerative Colitis. The progress in identifying genetic risk factors for UC has been slower than that for CD. This is mainly attributable to the fact that much of the attention for association studies was initially given to CD because it has a higher genetic contribution than UC.

However, early linkage analyses had been successful in identifying UC-specific linkage peaks, as well as highlighting the similarities between CD and UC. Interestingly, Giallourakis *et al* identified an IBD risk haplotype at the *IBD5* locus, which harbours a cytokine cluster, as a general IBD susceptibility factor, with an added epistatic effect between the *IBD5* risk haplotype and the *NOD2* mutations in the case of UC despite the fact that *NOD2* is not a risk factor for UC (Giallourakis et al., 2003). Additionally, the shared *IBD2* locus on chromosome 12 shows much stronger linkage for UC than for CD (Parkes et al., 2000).

The initial UC association analyses involved testing hits that were previously identified and confirmed for CD in UC datasets. Subsequently, genome-wide analysis of larger UC cohorts were carried out, where novel UC loci were identified, owing to the much larger sample sizes and hence higher power than could be envisaged just a few years before. These novel genes provide invaluable insight into the pathogenesis of UC. For instance, *CDH1*, *ECM* and *LAMB1* are all implicated in the correct functioning of the gut's mucosal barrier, which is at the forefront of IBD pathogenesis. In this context it is noteworthy that *CDH1*, encoding the E-Cadherin protein, was identified for the first time in this project as a susceptibility gene for CD, as is described in Chapter II of this thesis. In fact, *CDH1* is an interesting finding since it provides a direct link between UC and colorectal cancer, which is known to be more prevalent among UC cases. The promoter region of *CDH1* has been shown to be hypermethylated in colon cancer, which is an epigenetic mark commonly associated with gene silencing and thus, in this scenario, reduced expression of *CDH1*. This reduced expression could provide a tumour cell sufficient selective advantage for outgrowth (Wheeler et al., 2001). It has additionally been reported that a loss of E-Cadherin function displays preferential right-sided tumorigenesis as opposed to left-sided pathogenesis (Porter et al., 2002). Upon investigation of this left-side protection, it has been functionally shown that P-Cadherin, encoded by *CDH3* (See Chapter II), is expressed very early in colorectal dysplasia (Hardy et al., 2002) and an increase in P-Cadherin, which can substitute for many E-Cadherin functionalities, compensates for the down-regulation effect of E-Cadherin, namely decreased cell-adhesion and facilitation of cell migration (Porter et al., 2002).

Independently of *CDH1*, the risk of colorectal cancer in UC patients increases with disease duration as well as extent of disease, severity of inflammation and early onset of UC. This increased risk of colorectal cancer in UC patients could also in part be explained by the fact that premature telomere shortening, which is often associated with cancer, is observed in the colonic cells of UC patients (Risques et al., 2008).

The largest UC-specific GWAS carried out to date consisted of a meta-analysis of six GWASes containing approximately 6,687 cases and 19,718 controls as a discovery panel and 9,628 cases combined with 12,917 controls as a replication dataset. This study successfully increased the number of UC loci from 29 to 47. Subsequently, an IBD meta-analysis identified 12 new additional UC-specific loci. However, the total disease variance explained by these loci remains at 7.5%. The problem of missing heritability thus applies to UC as well.

Since the LDU mapping approach demonstrated a higher power to detect association and discover more susceptibility loci for CD, as described in Chapters II and III, the aim of the analysis that will be described in this chapter was to re-analyse the WTCCC Phase 2 UC dataset, followed by replication, in an attempt to unearth novel susceptibility loci for UC.

II. METHODS

i. Subjects

The UK WTCCC Phase II Ulcerative Colitis case-control dataset used in this part of the project consisted of 2,869 Ulcerative Colitis cases, and 5,986 nationally-ascertained controls prior to quality control. 2,988 control samples came from the National Blood Services (NBS) Collection of common controls and 2,998 control samples came from the 1958 British Birth Cohort (BC). The diagnosis of Ulcerative Colitis cases was done using standard endoscopic, radiologic and histopathological criteria and the samples were recruited from ten centres within the United Kingdom (Cambridge. Oxford, London, Newcastle, Sheffield, Edinburgh, Dundee, Manchester, Torbay and Exeter) (Consortium et al., 2009). The cases and the controls were genotyped on the high-resolution Affymetrix 6.0 array, containing approximately 1 million SNPs.

In order to replicate the findings, two replication datasets were used. Prior to quality control filtering, the first replication dataset consisted of 538 control samples, which came from a Parkinson's Disease GWA study and 487 Ulcerative Colitis cases came from a subset of the NIDDK International IBD Genetics Consortium UC data. All the NIDDK UC cases were reported to have the disease extending beyond the rectum (Silverberg et al., 2009). Both the cases and the controls for this replication dataset were genotyped on the Illumina 550v.3

genotyping platform consisting of approximately 550,000 SNPs. The samples for this case-control dataset were of white European, non-Hispanic ancestry (Silverberg et al., 2009).

The second replication dataset consisted of the other subset of the NIDDK UC cases containing 540 UC cases genotyped on the Illumina HumanHap 300v.2 genotyping platform. The ethnically-matched controls used were 515 Non-Jewish controls, which are a subset of the NIDDK CD GWAS described in Chapter II and III, genotyped on the Illumina HumanHap 300v.1 platform. Both the Illumina HumanHap 300v.1 and 300v.2 contain approximately of 320,000 SNPs. Different sets of control data were used since the NIDDK UC data did not have matching controls and therefore, ethnically-matched control data (PD GWAS Controls and NIDDK Non-Jewish CD GWAS Controls) genotyped on similar platforms were used in order to facilitate the merging of genotype data between the cases and controls. Table 11 summarises the data sets used in this part of the project, including the number of samples used after quality control, where the latter will be discussed in detail in the following section.

Table 11. Summary of datasets used for the UC GWAS.

Cases	Controls	Sample Size		Genotyping Platform
WTCCC2 UC	WTCCC2 NBS and 1958BC	Controls: Cases: **Total**	2,705 (1958BC) 2,674 (NBS) 2,361 **7,740**	Affymetrix 6.0 (~ 1 million SNPs)
NIDDK UC IBDO[§] UC GRU[§]	Parkinson's Disease GWAS Controls	Controls: Cases: **Total**	534 126 (UC IBDO) 359 (UC GRU) **1,019**	Illumina HumanHap550v.3 (~ 550,000 SNPs)
NIDDK UC GRU[§]	NIDDK CD Non-Jewish Controls	Controls: Cases: **Total**	503 534 **1,037**	*UC Cases:* Illumina HumanHap300v.2 *CD Controls:* Illumina HumanHap300v.1 (~ 320,000 SNPs)

[§] IBDO samples are patients that gave consent for their DNA to be used for IBD Research Only whereas GRU samples gave consent for their DNA to be used for General Research Use.

ii. Quality Control- WTCCC 2 UC Case-Control Dataset

The WTCCC provided quality control exclusion lists for both the SNP data as well as the sample data for each of the UC, NBS and 1958BC data-sets. Before carrying out any quality control analyses myself, the data were filtered according to these lists. The SNP exclusion list contained 5 exclusion criteria for each SNP, namely Minor Allele Frequency (MAF), information, Hardy Weinberg Equilibrium (HWE) test, missingness and plate effect. If a SNP failed to pass any one of these criteria, it was advised that the SNP should be excluded from the data. The exclusion thresholds that the WTCCC used for each of these criteria will be explained in more detail below.

For this WTCCC SNP exclusion list, the MAF threshold used for each SNP was 1%. Therefore, any SNP where the MAF falls below 1% was included in the exclusion list. The information criterion is a measure of statistical information about allele frequency in the genotype data, ranging from 0 to 1 and reflects the level of certainty regarding the called genotype, where 0 denotes complete uncertainty and 1 complete certainty. In order for SNPs to be retained in the data, the information had to be higher that 0.975. The HWE test is a measure of deviation from HWE in the general population. This test is generally carried out in controls since true association in cases can distort HWE tests. Since in most cases deviation from HWE generally reflects poor genotyping, or copy number variation for which the analysis would be confusing, HWE testing is routinely done for quality control prior to association studies. The P-value threshold used

by the WTCCC for the HWE test for each SNP was $1x10^{-20}$ and therefore, any SNPs with a lower P-value was incorporated in the exclusion list. The missingness relates to the proportion of genotypes, which fail to reach the maximum call probability threshold of 0.90. If this missingness proportion is higher than 0.02, the SNP is marked for exclusion. The plate criterion is introduced because cases and controls were plated on different plates and consists of a P-value of a *n*-degree of freedom test of plate association. If the P-value is lower than $1x10^{-05}$, the SNP is marked for exclusion.

The WTCCC sample exclusion list considered six criteria for quality control, namely SNP missingness across a sample, relatedness of samples, ancestry, gender, channel and identity. If a sample failed to pass any one of these criteria, it was suggested that the sample should be excluded from the data.

The missingness criterion can be subdivided into two sub-criteria. Firstly, the individuals were compared on the basis of a fraction of successfully called SNPs that were heterozygous, and then on the proportion of SNPs where the maximum call probability did not meet the 0.90 threshold. If any individuals were outlying in either sub-criteria, they were listed in the exclusion list. The individuals in the data were tested for relatedness by the WTCCC by running a Hidden Markov Model (HMM), which identifies the proportion Identity-by-Descent (IBD) between pairs of individuals. In this analysis, if two individuals were more than 5% identical by descent, the individual with the lowest genotype call rate was marked for exclusion from the data. For the ancestry analysis, the individuals were projected onto the first two principal components of a Principal Components Analysis (PCA)

of HapMap individuals. If any individuals were shown to differ from the remaining individuals in the dataset in terms of ancestry, they were flagged for exclusion in the WTCCC exclusion list provided. The WTCCC inferred the gender from the intensities of the A-allele probes on the X chromosome, which were normalized against the intensities of the autosomal SNPs. If certain individuals showed discrepancies between the gender supplied and the inferred gender, or if the gender could not be inferred from the intensities, these individuals were listed in the exclusion list. The channel criterion consisted of excluding individuals where the mean of their A and B allele intensities from 10,000 SNPs on chromosome 22 was outlying when compared to the sample at large. The identity analysis that the WTCCC carried out checked the genotype concordance of approximately 30 SNPs that were previously genotyped using the Sequenom at the Wellcome Trust Sanger Institute (WTSI) and the genotypes obtained for the Affymetrix 6.0 platform for these SNPs for each individual. If the concordance between the two platforms was lower than 90%, the individual was marked for exclusion from the data.

After filtering the data for these SNP and individual exclusion criteria, the routine and widely-applied GWAS quality control outlined in the next section was carried out. Appendix III and IV illustrate flowcharts outlining the steps carried out for the data manipulation of the WTCCC2 and NIDDK Data.

a. Principal Components Analysis

After removing SNPs and individuals that were indicated in the sample exclusion lists provided by the WTCCC, Principal Components Analysis (PCA) was carried out to detect possible outliers in the dataset. In order to perform the PCA, any SNPs with a Minor Allele Frequency (MAF) smaller than 0.05, missing rate of more than 5% and an LD r^2 threshold larger than 0.2 were removed. The LD pruning was done in order to avoid a heavy influence from clusters of SNPs in LD with each other for both the PCA and Relatedness analysis described below. The PCA was carried out in the publically available software R using the packages gdsfmt and SNPRelate (Zheng et al., 2012). The first step was to calculate the genetic covariance matrix from the genotype data, as well as the correlation coefficients between the sample loadings and the SNP genotypes, followed by the calculation of SNP eigenvectors (Zheng et al., 2012). The PCA results, as illustrated in Figure 12, showed that there were two outliers, one in each of the 1958BC and NBS datasets. Despite previous filtering for identity by descent outliers using the WTCCC exclusion lists, subsequent analysis of these two individuals showed that they were almost certainly the same individual or identical twins, since visual inspection of the data showed that the SNP genotypes were the same for both individuals, with the only difference in the genotypes due to missingness. This could be due to the fact that separate WTCCC exclusion lists were provided for the two control groups (1958BC and NBS) thus suggesting that QC was carried out on the two data separately, which would have missed identifying these two related individuals. Therefore, the sample that had most missing genotypes was removed from the data before further proceeding with the quality control analysis.

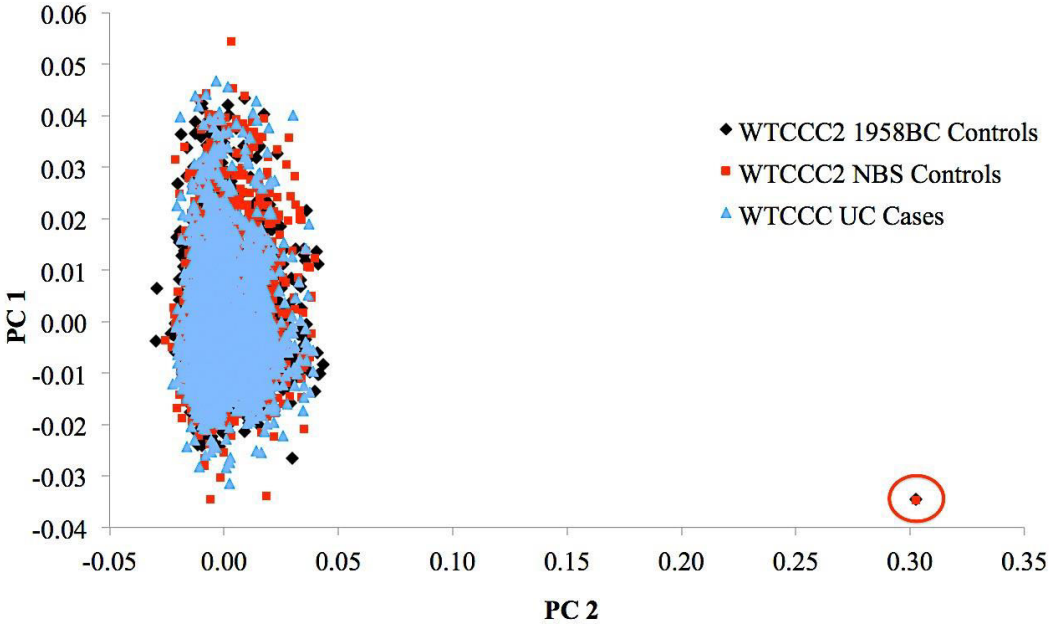


Figure 12. Principal Components Analysis of WTCCC Phase II Case-Control Dataset.
The Y-axis represents the first Principal Component and the X-axis represents the second Principal Component. The blue triangles represent the sample loadings for the WTCCC UC Cases, the black diamonds the sample loadings for the WTCCC 1958BC Controls and the red squares the sample loadings for the WTCCC NBS controls. The samples circled in red are the outliers identified in this analysis and the outlier with most missingness was removed from the data.

The PCA procedure was repeated using the WTCCC2 UC Case-Control Data merged with the HapMap 3 dataset. This was done in order to see how the WTCCC2 UC Case-Control dataset clusters with respect to the populations in HapMap 3 (See Figure 13). Since the two datasets came from different studies, the strand orientation was not the same for all SNPs. To overcome this problem, the data were merged in PLINK, which flags any SNPs found to be in opposite strand orientation between both datasets, followed by flipping the strand for the

SNPs in question in the WTCCC2 UC Case-Control data to match that in HapMap3. Additionally, since the Affymetrix 6.0 genotyping platform contained palindromic SNPs (A/T or G/C SNPs), these needed to be removed from the data for the quality control analysis only, prior to merging with the HapMap 3 data as the strand orientation of these SNPs between two different datasets cannot be determined.

As expected, the data cluster with the European population CEU, which consists of Utah residents with Northern and Western European ancestry from the CEPH (Centre d'Étude du Polymorphisme Humain) collection and with the Tuscan population from Italy, (TSI).

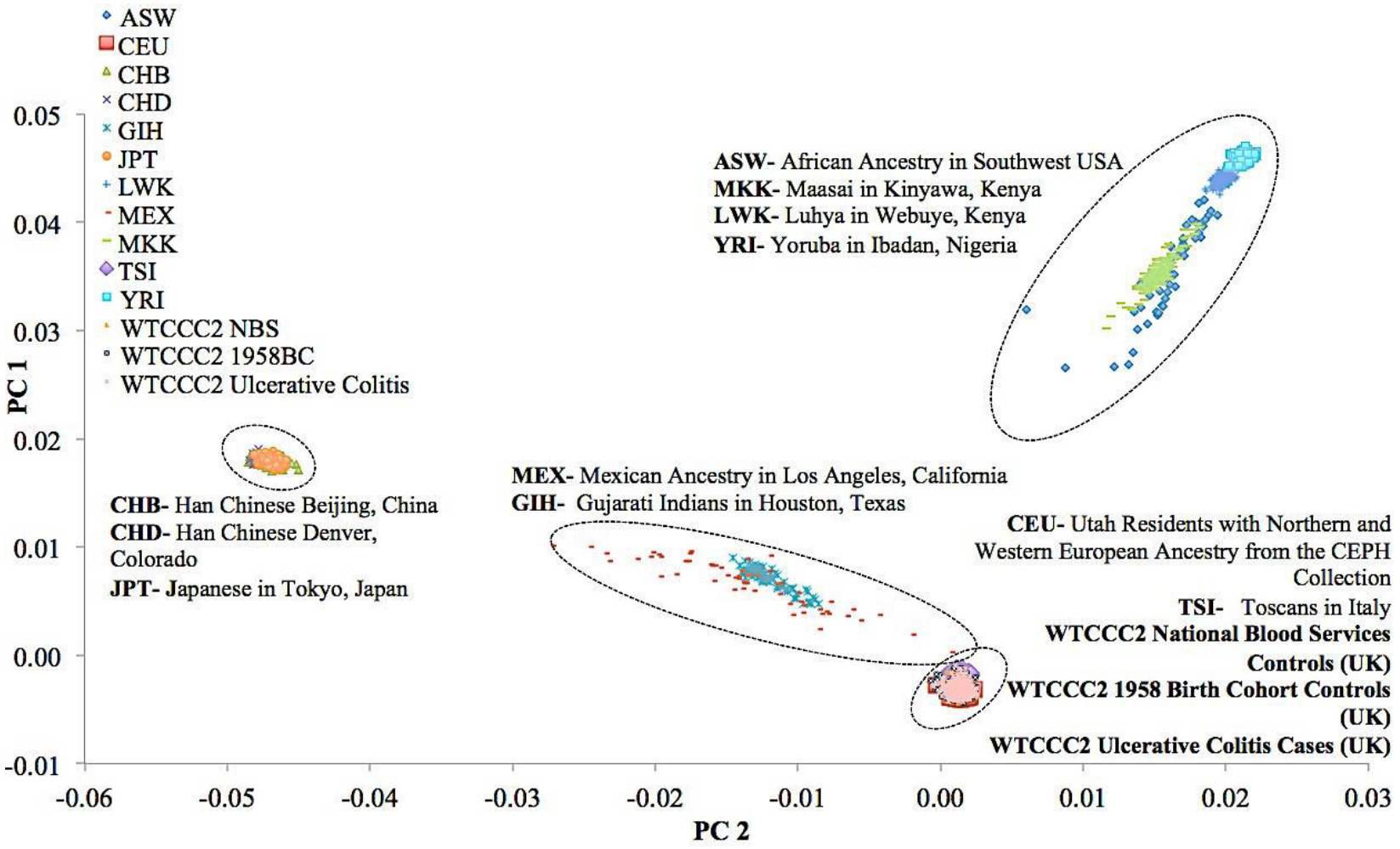

ASW
CEU
CHB
CHD
GIH
JPT
LWK
MEX
MKK
TSI
YRI
WTCCC2 NBS
WTCCC2 1958BC
WTCCC2 Ulcerative Colitis
ASW- African Ancestry in Southwest USA
MKK- Maasai in Kinyawa, Kenya
LWK- Luhya in Webuye, Kenya
YRI- Yoruba in Ibadan, Nigeria
CHB- Han Chinese Beijing, China
CHD- Han Chinese Denver, Colorado
JPT- Japanese in Tokyo, Japan
MEX- Mexican Ancestry in Los Angeles, California
GIH- Gujarati Indians in Houston, Texas
CEU- Utah Residents with Northern and Western European Ancestry from the CEPH Collection
TSI- Toscans in Italy
WTCCC2 National Blood Services Controls (UK)
WTCCC2 1958 Birth Cohort Controls (UK)
WTCCC2 Ulcerative Colitis Cases (UK)
PC 1
PC 2
0.05
0.04
0.03
0.02
0.01
0.00
-0.01
-0.06
-0.05
-0.04
-0.03
-0.02
-0.01
0.00
0.01
0.02
0.03

Figure 13. Principal Components Analysis of merged HapMap3 and WTCCC2 UC Case-Control Datasets.
The Y-axis represents the first Principal Component and the X-axis represents the second Principal Component. The data points on the plot illustrate the sample loadings. The WTCCC Case-Control data cluster with the CEU and TSI, as expected.

b. Relatedness Analysis- Identity-by-Descent and Identity-by-State Analysis

The relatedness analysis was equally carried out in R using the packages gdsfmt and SNPRelate.

The Identity-by-Descent (IBD) analysis using the PLINK Method of Moments (MoM) was carried out in order to identify whether any individuals in the UC case-control data set were related after the initial individual filtering. As expected, since the related individuals had already been removed, the results of this analysis showed no significant relatedness between the individuals, as shown in Figure 14.

The Identity-by-State (IBS) analysis was done by creating a $n \times n$ matrix containing genome-wide average IBS pairwise identities, followed by a Multidimensional Scaling (MDS) analysis, which allowed the visualisation of similarities and differences in the individuals under study. The IBS and MDS analyses were carried out on the HapMap3 dataset merged with the WTCCC2 UC Case-Control data, as was done for the PCA described in the previous section. The results of the MDS analysis, as illustrated in Figure 15, demonstrate that no outliers were detected and also that allele sharing was, as expected, highest between the WTCCC UC Case-Control dataset, the CEU and TSI since all three are European populations.

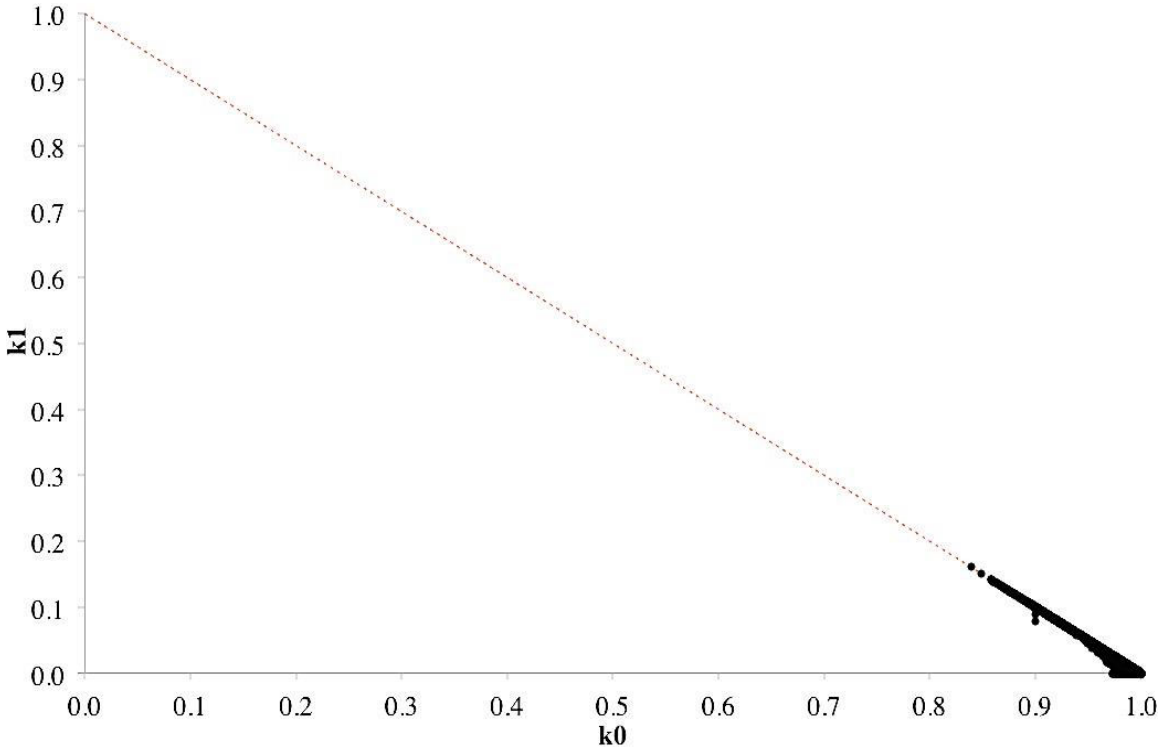


Figure 14. Identity-by-Descent Analysis of the merged WTCCC2 UC Case-Control Dataset based on pairwise kinship coefficients.
The Y-axis represents the k1 pairwise IBD coefficients, which are probabilities of sharing one IBD. The X-axis represents the k0 pairwise IBD coefficients, which is the probability of sharing zero IBD. The red dashed line represents the expected IBD. The black circles represent the samples in the WTCCC UC Case-Control data. The results show that there is no relatedness among the samples in the WTCCC Case-Control data.

c. SNP filtering and individual missingness

In order to minimize false positive associations due to artifacts in the data the quality control was made more stringent. SNPs with a MAF lower than 0.03, HWE test P-value lower than 1.57×10^{-3} and a missing rate greater than 20% were further excluded from the data. The final sample count is listed in Table 11.

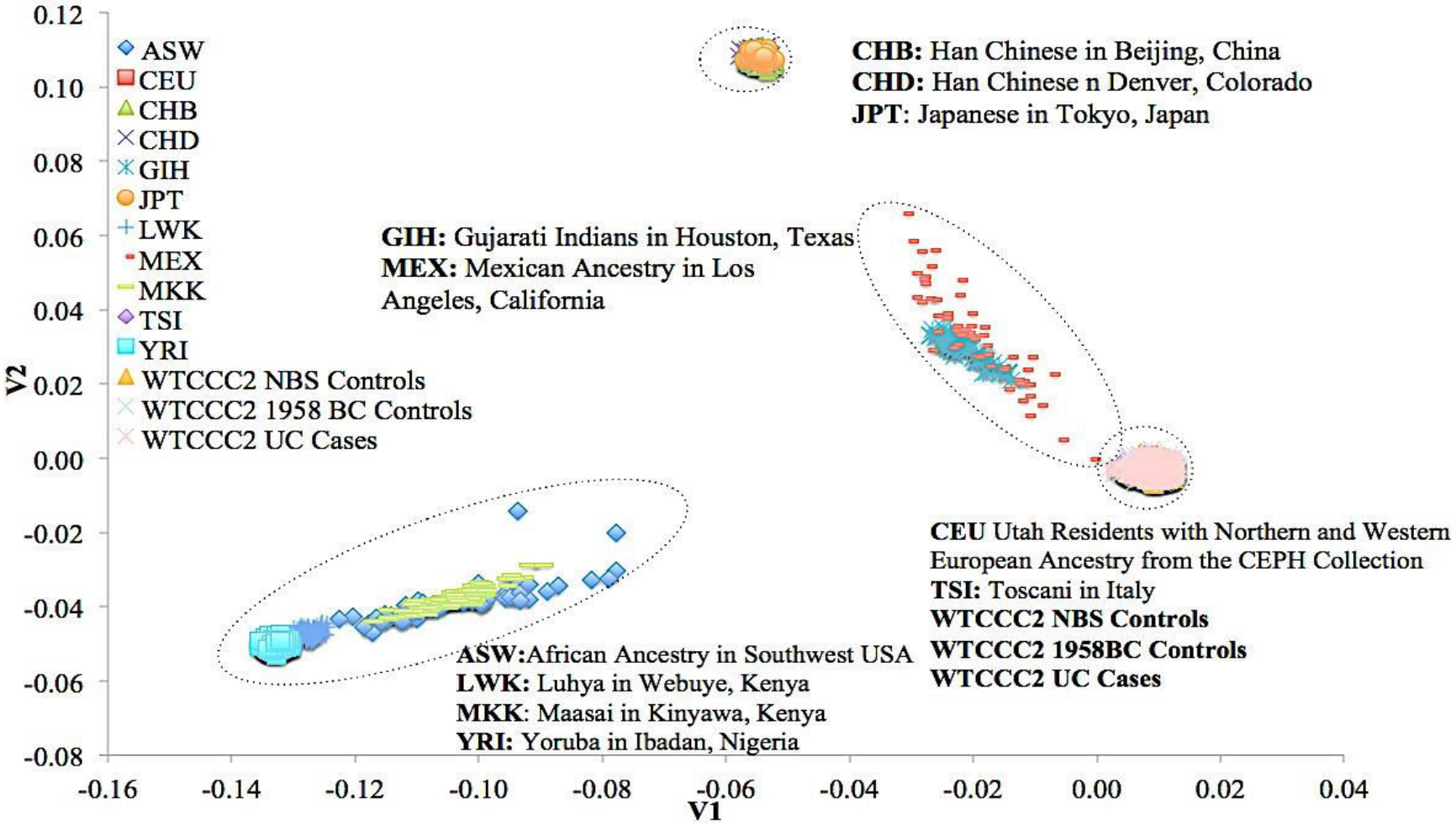

0.12
0.10
0.08
0.06
0.04
0.02
0.00
-0.02
-0.04
-0.06
-0.08
V2
-0.16
-0.14
-0.12
-0.10
-0.08
-0.06
-0.04
-0.02
0.00
0.02
0.04
V1
ASW
CEU
CHB
CHD
GIH
JPT
LWK
MEX
MKK
TSI
YRI
WTCCC2 NBS Controls
WTCCC2 1958 BC Controls
WTCCC2 UC Cases
CHB: Han Chinese in Beijing, China
CHD: Han Chinese n Denver, Colorado
JPT: Japanese in Tokyo, Japan
GIH: Gujarati Indians in Houston, Texas
MEX: Mexican Ancestry in Los Angeles, California
CEU Utah Residents with Northern and Western European Ancestry from the CEPH Collection
TSI: Toscani in Italy
WTCCC2 NBS Controls
WTCCC2 1958BC Controls
WTCCC2 UC Cases
ASW:African Ancestry in Southwest USA
LWK: Luhya in Webuye, Kenya
MKK: Maasai in Kinyawa, Kenya
YRI: Yoruba in Ibadan, Nigeria

Figure 15. Multidimensional Scaling Analysis, basd on IBS distances of the merged HapMap3 and WTCCC2 UC Case-Control dataset.
The Y-axis represents the second vector of the MDS and the X-axis the first vector. The data points on the plot represent the HapMap3 and WTCCC Case-Control samples. No outliers were detected in this analysis.

iii. Quality Control Analysis- NIDDK UC datasets

For the additional two Ulcerative Colitis Datasets, the same quality control procedure outlined in section II.ii.a-c of this Chapter was carried out. However, prior to combining the UC Cases with the Controls, the PCA was carried out on the separate datasets, since no prior exclusion lists were given for the separate data unlike for the WTCCC2 UC described before. The NIDDK UC IBDO Cases (See Table 11) showed no outliers (See Figure 16) but GRU Cases genotyped on the Illumina HumanHap550v.3 (See Table 11) showed 1 major outlier, as shown circled in Figure 17. This was expected, as in the NIDDK data release notes it was noted that after the original analysis of this data it was subsequently identified that one of the samples was Trinidadian. Similarly, the Parkinson's Disease Control data also showed four outliers (See Figure 18).

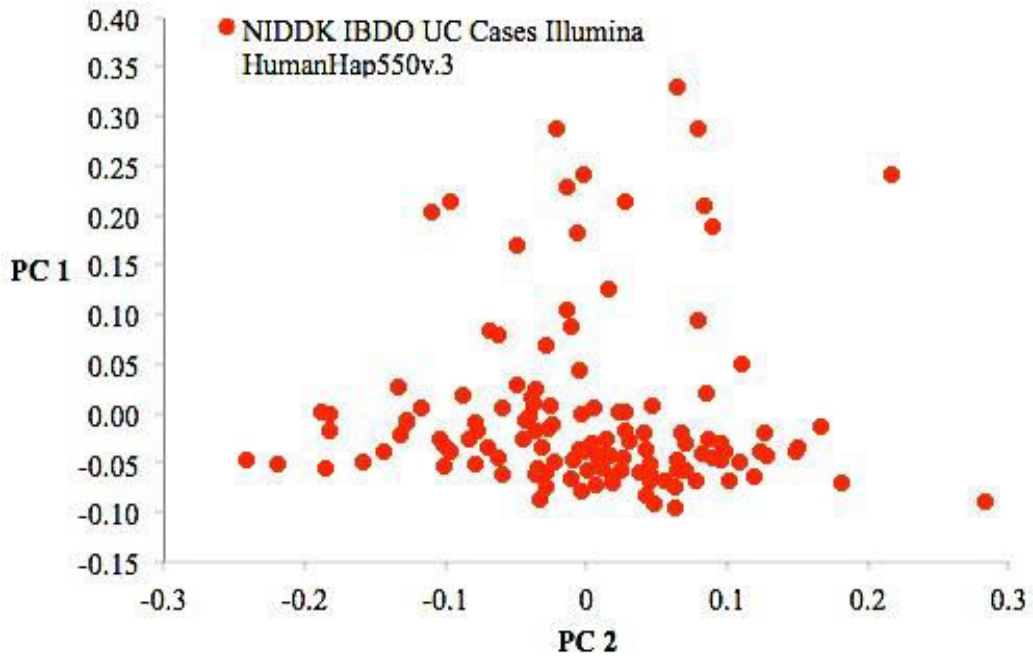


Figure 16. Principal Components Analysis of NIDDK IBDO UC Cases genotyped on the Illumina HumanHap 550v.3 platform.
The Y-axis represents the first Principal Component and the X-axis represents the second Principal Component. The red circles represent the sample loadings for the NIDDK IBDO UC Cases. No outliers were detected in this analysis. The NIDDK IBDO UC Cases are patients who gave consent for their DNA to be used for IBD research only.

After removing the outliers from both datasets, the Parkinson's Disease GWAS controls were merged with the NIDDK UC IBDO and GRU cases genotyped on the Illumina HumanHap550v3. Again, since the two datasets came from different studies, there were differences in strand orientation for certain SNPs, despite being genotyped on the same platform. Therefore, in order to merge the two datasets, the same procedure in PLINK was followed as described earlier. However, in this case, the Illumina HumanHap550v.3 did not contain any palindromic SNPs. After the Cases and Controls were successfully merged, the

Case-Control dataset was then merged with the HapMap 3 data, after flipping the strand orientation to match that of HapMap 3, and the PCA was carried out (Figure 19).

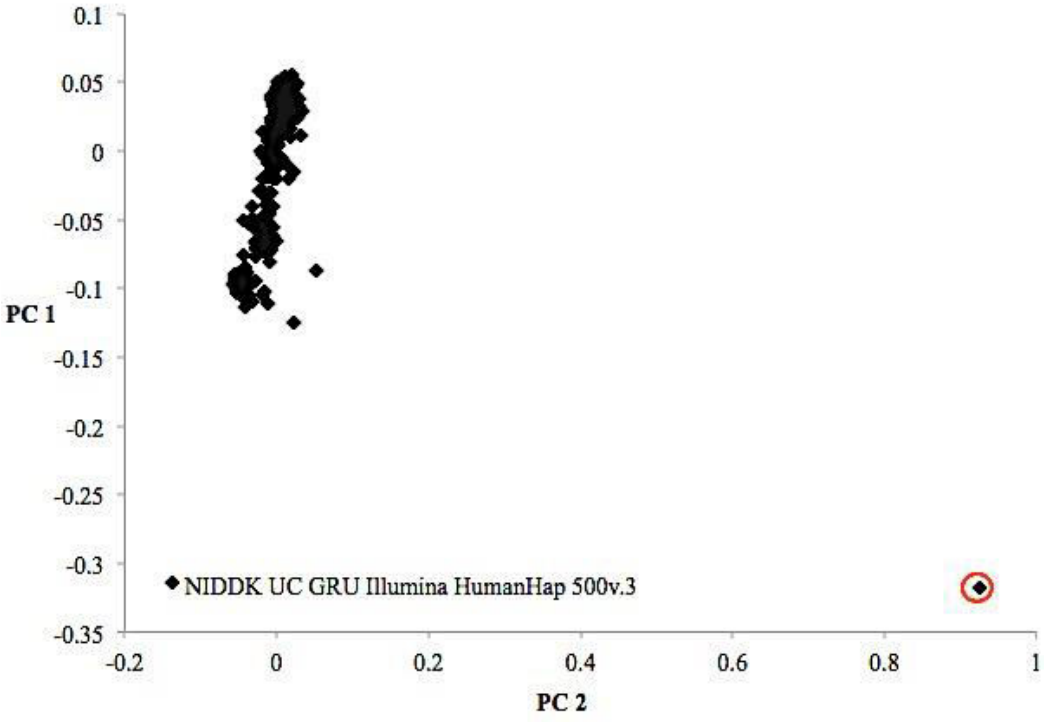


Figure 17. Principal Components Analysis of NIDDK GRU UC Cases genotyped on the Illumina HumanHap 550v.3 platform.
The Y-axis represents the first Principal Component and the X-axis represents the second Principal Component. The black diamonds represent the sample loadings for the NIDDK GRU UC Cases. The samples circled in red represents the outlier identified in this analysis, which was removed from the data. The NIDDK GRU UC Cases are patients who gave consent for their DNA to be used for general research use.

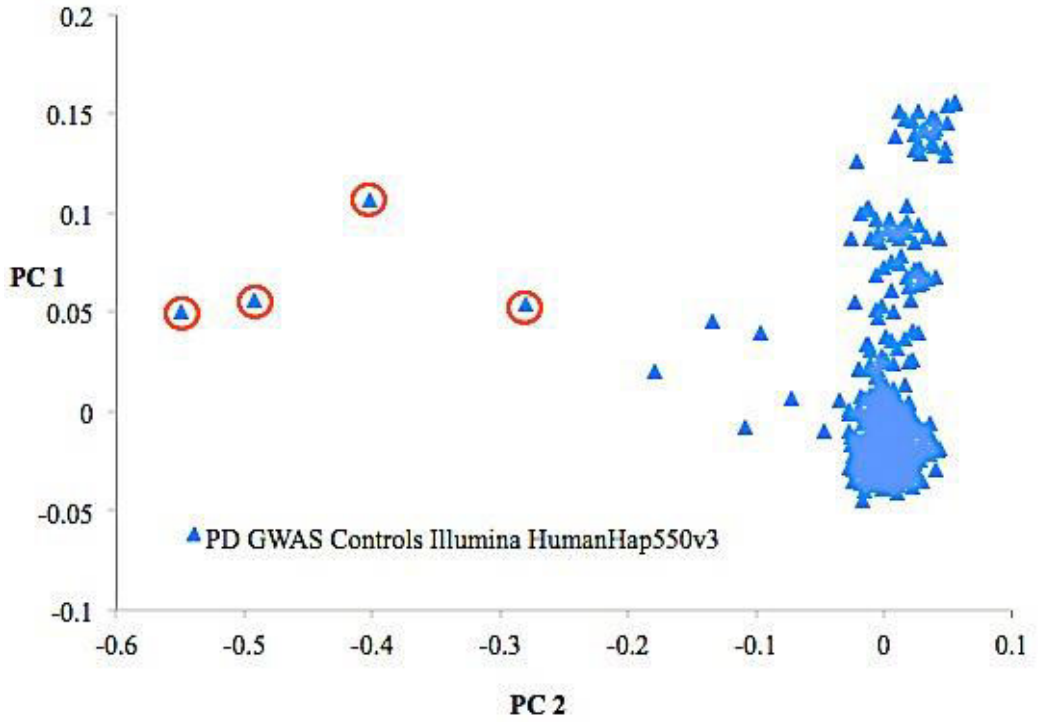


Figure 18. Principal Components Analysis of Parkinson's Disease GWAS Control data genotyped on the Illumina HumanHap 550v.3 platform.
The Y-axis represents the first Principal Component and the X-axis represents the second Principal Component. The blue triangles represent the sample loadings for the PD GWAS Controls. The samples circled in red are the outliers identified in this analysis, which were removed from the data.

Despite the previous filtering, additional outliers could still be detected when the NIDDK IBDO and GRU UC Cases genotyped on the Illumina HumanHap 550v.3 were merged with the PD GWAS Control data and with the HapMap3 data (Figure 19). Two outliers were found in the NIDDK UC GRU cases and one in the Parkinson's Disease GWAS controls. These three outliers were also detected in the IBS-based MDS analysis (Figure 20) but no outliers were visible in the IBD analysis (Figure 21).

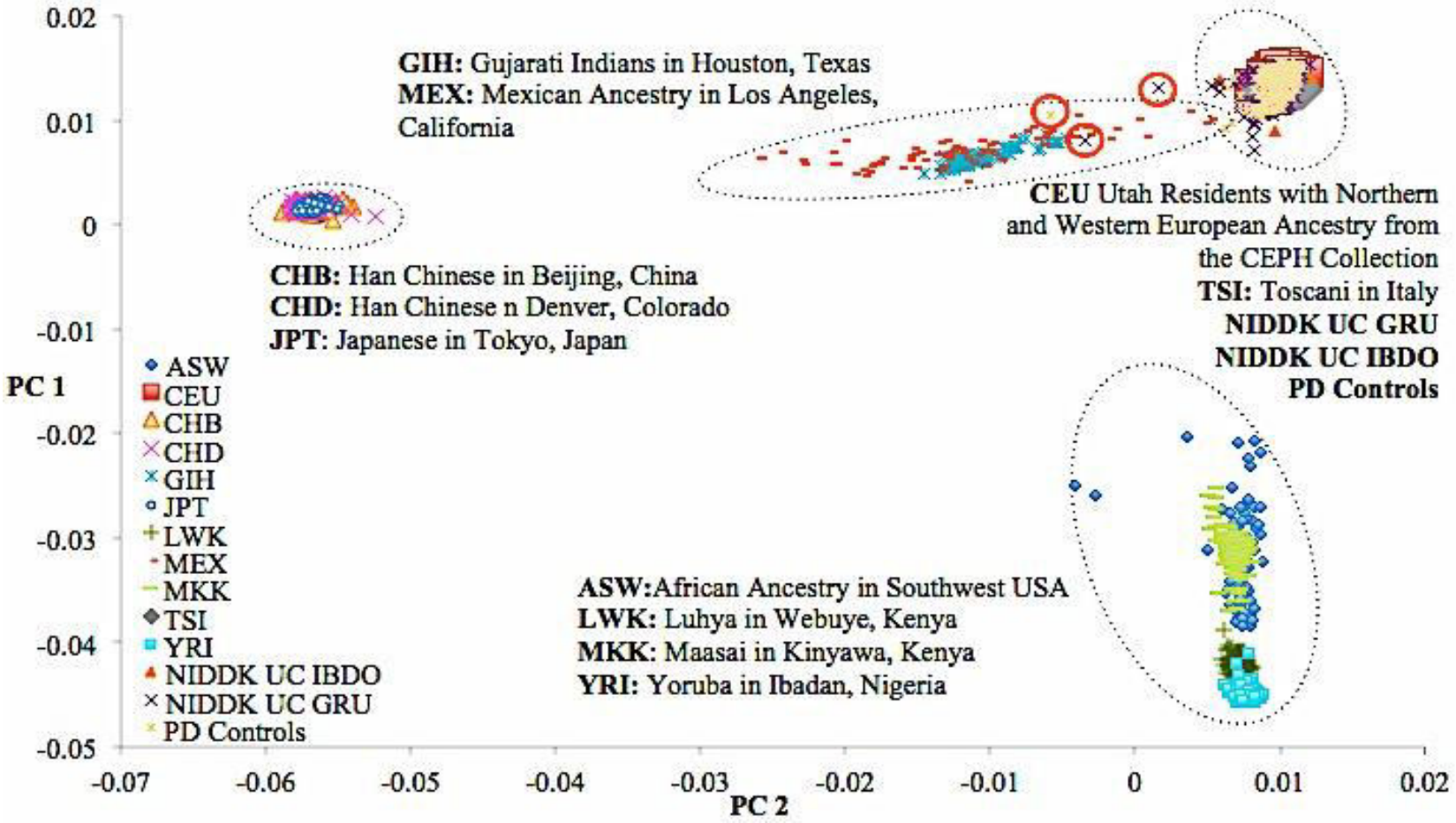

0.02
0.01
0
-0.01
-0.02
-0.03
-0.04
-0.05
PC 1
-0.07
-0.06
-0.05
-0.04
-0.03
-0.02
-0.01
0
0.01
0.02
PC 2
ASW
CEU
CHB
CHD
GIH
JPT
LWK
MEX
MKK
TSI
YRI
NIDDK UC IBDO
NIDDK UC GRU
PD Controls
GIH: Gujarati Indians in Houston, Texas
MEX: Mexican Ancestry in Los Angeles, California
CEU Utah Residents with Northern and Western European Ancestry from the CEPH Collection
TSI: Toscani in Italy
NIDDK UC GRU
NIDDK UC IBDO
PD Controls
CHB: Han Chinese in Beijing, China
CHD: Han Chinese n Denver, Colorado
JPT: Japanese in Tokyo, Japan
ASW:African Ancestry in Southwest USA
LWK: Luhya in Webuye, Kenya
MKK: Maasai in Kinyawa, Kenya
YRI: Yoruba in Ibadan, Nigeria

Figure 19. Principal Compoenents Analysis of NIDDK IBDO and GRU UC Cases merged with Parkinson's Disease GWAS Controls genotyped on the Illumina HumanHap 550v.3 and HapMap3 data.
The Y-axis represents the first Principal Component and the X-axis represents the second Principal Component. The data points on the plot represent the sample loadings for the HapMap3, PD GWAS Controls and NIDDK UC data. The samples circled in red are the outliers identified in this analysis, which were removed from the data.

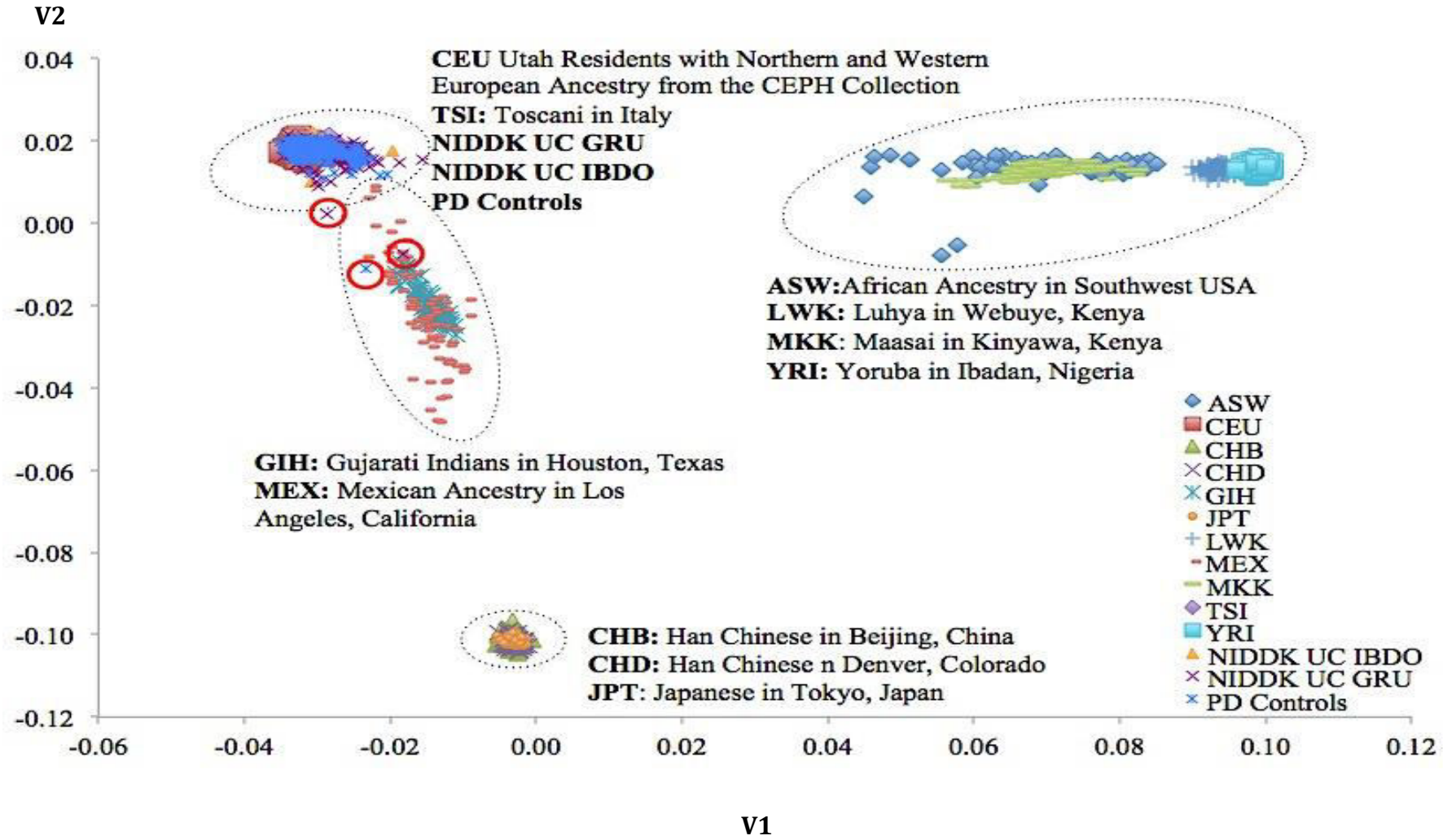

V2
0.04
0.02
0.00
-0.02
-0.04
-0.06
-0.08
-0.10
-0.12
-0.06
-0.04
-0.02
0.00
0.02
0.04
0.06
0.08
0.10
0.12
V1
CEU Utah Residents with Northern and Western European Ancestry from the CEPH Collection
TSI: Toscani in Italy
NIDDK UC GRU
NIDDK UC IBDO
PD Controls
ASW:African Ancestry in Southwest USA
LWK: Luhya in Webuye, Kenya
MKK: Maasai in Kinyawa, Kenya
YRI: Yoruba in Ibadan, Nigeria
GIH: Gujarati Indians in Houston, Texas
MEX: Mexican Ancestry in Los Angeles, California
CHB: Han Chinese in Beijing, China
CHD: Han Chinese n Denver, Colorado
JPT: Japanese in Tokyo, Japan
ASW
CEU
CHB
CHD
GIH
JPT
LWK
MEX
MKK
TSI
YRI
NIDDK UC IBDO
NIDDK UC GRU
PD Controls

Figure 20. Multidimensional Scaling Analysis based on IBS distances of NIDDK IBDO and GRU UC Cases merged with Parkinson's Disease GWAS Controls genotyped on the Illumina HumanHap 550v.3 platform and HapMap3 data.
The Y-axis represents the second vector of the MDS and the X-axis the first vector. The data points on the plot represent the HapMap3, the PD GWAS controls and NIDDK UC Data. The samples circled in red represent the outliers identified in this analysis, which are the same as those illustrated in Figure 19. The samples were removed from the data.

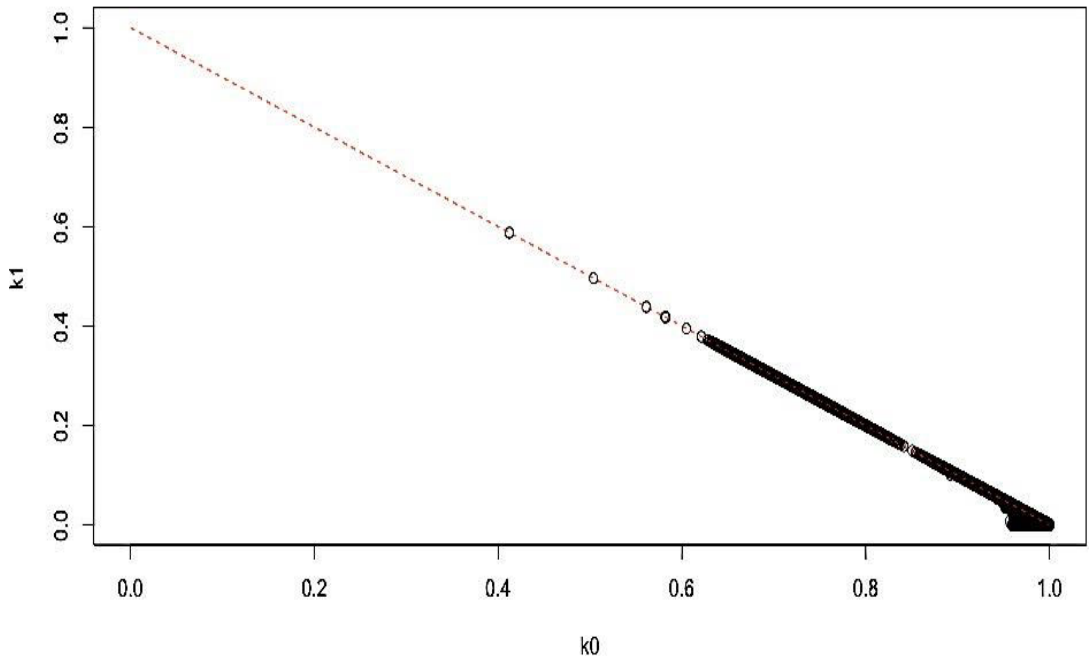


Figure 21. IBD analysis of the merged Parkinson's Disease GWAS Controls and NIDDK IBDO and GRU UC Cases genotyped on the Illumina HumanHap 500v.3 platform based on pairwise kinship coefficients.
The Y-axis represents the k1 pairwise IBD coefficients, which are probabilities of sharing one IBD. The X-axis represents the k0 pairwise IBD coefficients, which is the probability of sharing zero IBD. The red dashed line represents the expected IBD. The black circles represent the samples in the PD GWAS controls and NIDDK IBDO and GRU UC Cases. The results show that there is no relatedness among the samples analysed.

The final step in the quality control filtering of this case-control dataset, SNPs with a MAF higher than 0.03, HWE test P-value lower than 1.57×10^{-3} and a missing rate greater than 20%. The final SNP and individual count is listed in Table 11.

The same quality control procedure was carried out for the remaining NIDDK GRU UC Cases, which were genotyped on the Illumina HumanHap300v.2 platfrom and the NIDDK Non-Jewish CD Controls dataset, genotyped on the

Illumina HumanHap300v.1 (See Table 11). Although both datasets were genotyped on the Illumina HumanHap300, they were genotyped on different versions of the platform and therefore they did not contain the same set of SNPs. Therefore, prior to merging the NIDDK UC GRU Cases and the NIDDK Non-Jewish CD Controls, the SNPs that were not present on both platforms were removed. Subsequently, merging of the dataset was carried out in PLINK, as described earlier, after the necessary strand orientation flipping was done. After the successful merging of the cases with the controls, the quality controls analysis described in II.ii.a-c was carried out.

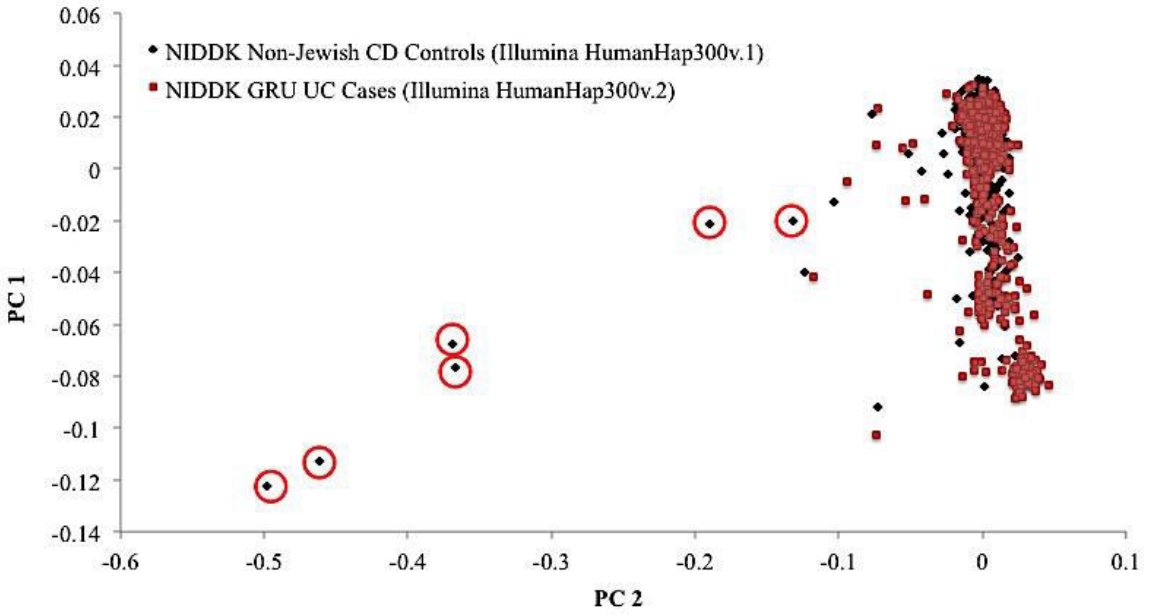


Figure 22. Principal Components Analysis of merged NIDDK Non-Jewish CD Controls genotyped on the Illumina HumanHap 300v.1 and NIDDK GRU UC Cases genotyped on the Illumina HumanHap 300v.2 platform.
The Y-axis represents the first Principal Component and the X-axis represents the second Principal Component. The red squares represent the sample loadings for the NIDDK GRU UC Cases and the black diamonds represent the sample loadings for the NIDDK Non Jewish CD Controls. The samples circled in red are the outliers identified in this analysis, which were removed from the data.

The PCA revealed six outliers in the data, all of which came from the NIDDK Non-Jewish Control data, as shown circled in red in Figure 22. After these six outliers were removed from the data and the data were merged with the HapMap3 data, five outliers in the NIDDK GRU UC cases and four outliers in the NIDDK Non Jewish CD Controls were detected in a PCA analysis (Figure 23), as well as in the MDS analyses based on IBS distances (Figure 24). These individuals were removed from the data prior to proceeding with the quality control analysis.

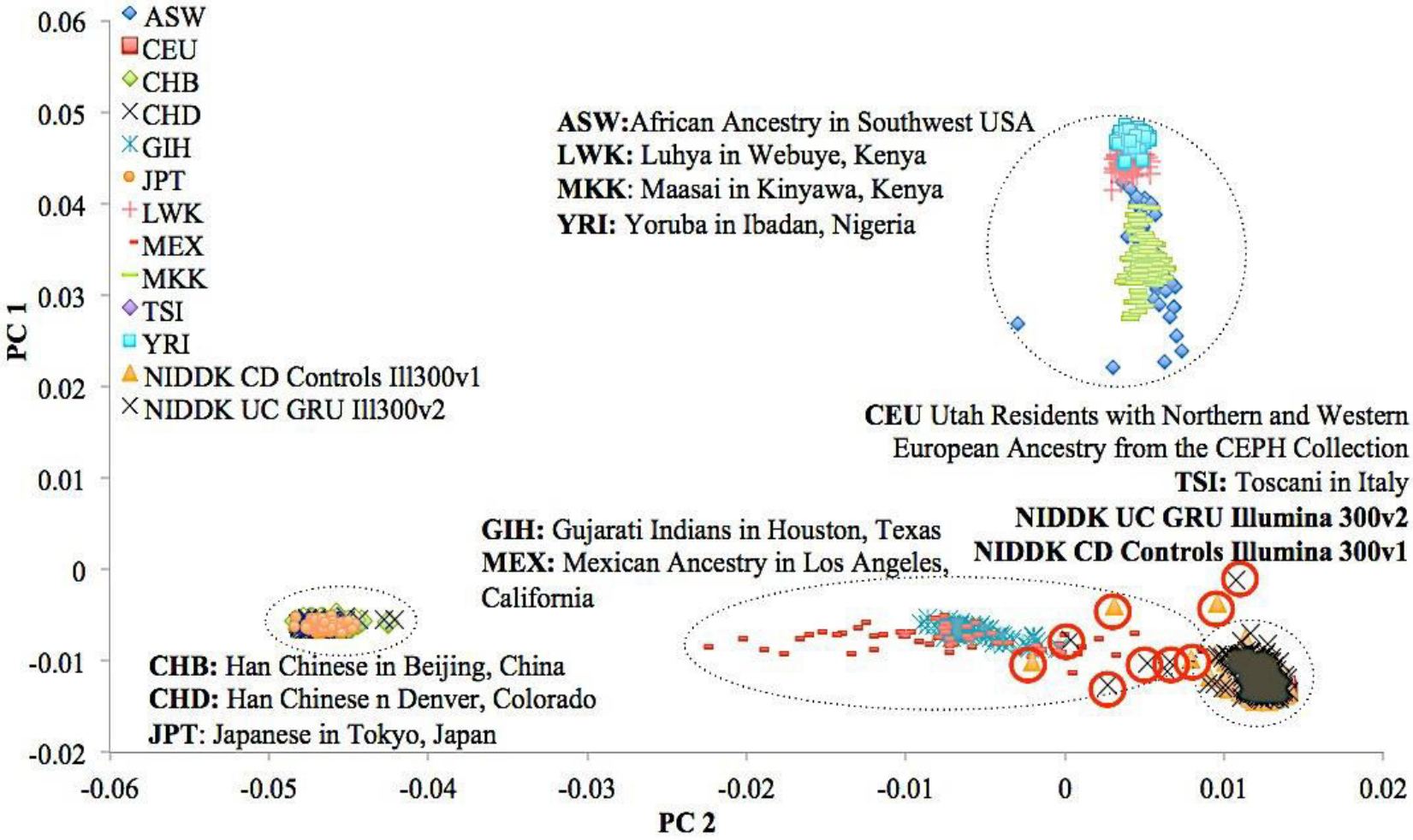

ASW
CEU
CHB
CHD
GIH
JPT
LWK
MEX
MKK
TSI
YRI
NIDDK CD Controls Ill300v1
NIDDK UC GRU Ill300v2
ASW:African Ancestry in Southwest USA
LWK: Luhya in Webuye, Kenya
MKK: Maasai in Kinyawa, Kenya
YRI: Yoruba in Ibadan, Nigeria
CEU Utah Residents with Northern and Western European Ancestry from the CEPH Collection
TSI: Toscani in Italy
NIDDK UC GRU Illumina 300v2
NIDDK CD Controls Illumina 300v1
GIH: Gujarati Indians in Houston, Texas
MEX: Mexican Ancestry in Los Angeles, California
CHB: Han Chinese in Beijing, China
CHD: Han Chinese n Denver, Colorado
JPT: Japanese in Tokyo, Japan
0.06
0.05
0.04
0.03
0.02
0.01
0
-0.01
-0.02
PC 1
-0.06
-0.05
-0.04
-0.03
-0.02
-0.01
0
0.01
0.02
PC 2

Figure 23. Principal Components Anlaysis of merged NIDDK Non-Jewish CD Controls (Illumina HumanHap 300v.1), NIDDK GRU UC Cases (Illumina HumanHap 300v.2) and HapMap3 data.

The Y-axis represents the first Principal Component and the X-axis represents the second Principal Component. The data points on the plot represent the sample loadings for the HapMap3, NIDDK Non Jewish CD Controls and NIDDK GRU UC Cases. The samples circled in red are the outliers identified in this analysis, which were removed from the data.

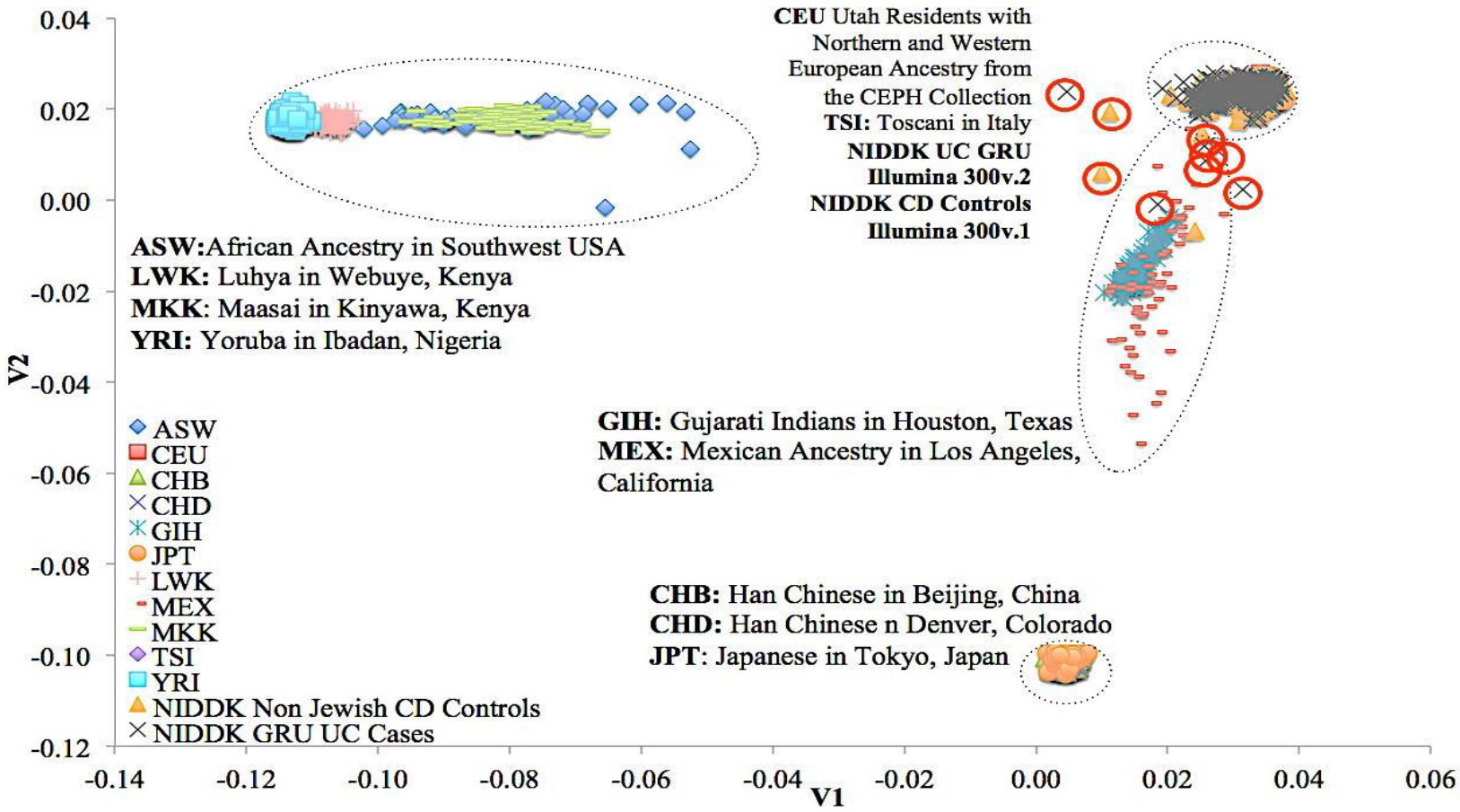

0.04
0.02
0.00
-0.02
-0.04
-0.06
-0.08
-0.10
-0.12
V2
-0.14
-0.12
-0.10
-0.08
-0.06
-0.04
-0.02
0.00
0.02
0.04
0.06
V1
ASW:African Ancestry in Southwest USA
LWK: Luhya in Webuye, Kenya
MKK: Maasai in Kinyawa, Kenya
YRI: Yoruba in Ibadan, Nigeria
CEU Utah Residents with Northern and Western European Ancestry from the CEPH Collection
TSI: Toscani in Italy
NIDDK UC GRU Illumina 300v.2
NIDDK CD Controls Illumina 300v.1
GIH: Gujarati Indians in Houston, Texas
MEX: Mexican Ancestry in Los Angeles, California
CHB: Han Chinese in Beijing, China
CHD: Han Chinese n Denver, Colorado
JPT: Japanese in Tokyo, Japan
ASW
CEU
CHB
CHD
GIH
JPT
LWK
MEX
MKK
TSI
YRI
NIDDK Non Jewish CD Controls
NIDDK GRU UC Cases

Figure 24. Multidimensional Scaling Analysis of merged NIDDK Non-Jewish CD Controls, NIDDK GRU UC Cases and HapMap3 data, basd on IBS Distances.
The Y-axis represents the second vector of the MDS and the X-axis the first vector. The data points on the plot represent the HapMap3, the NIDDK Non Jewish CD Controls and NIDDK GRU UC Data. The samples circled in red represent the outliers identified in this analysis, which are the same as those illustrated in Figure 19. The samples were removed from the data.

The subsequent IBD analysis also revealed two pairs of related individuals (Figure 25). The first pair consisted of two Non-Jewish CD controls whereas the second pair consisted of a Non-Jewish CD control and a GRU UC Case. One member of the first pair was removed and the NIDDK Non Jewish CD control was removed from the second pair.

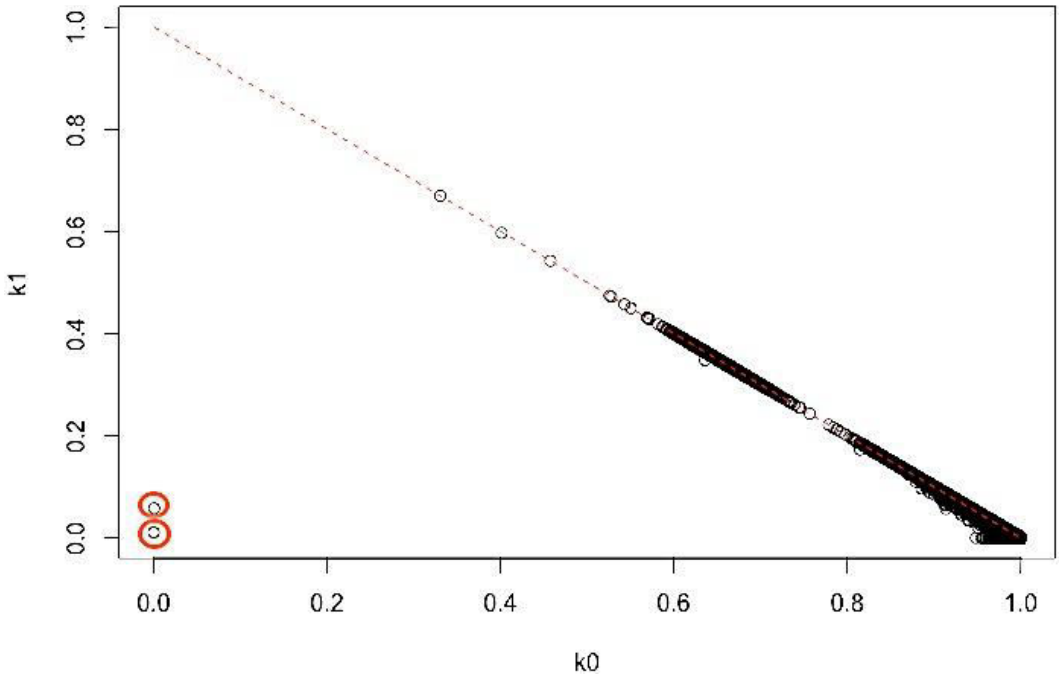


Figure 25. IBD analysis of the merged NIDDK Non-Jewish CD Controls and NIDDK GRU UC Cases based on pairwise kinship coefficients.
The Y-axis represents the k1 pairwise IBD coefficients, which are probabilities of sharing one IBD. The X-axis represents the k0 pairwise IBD coefficients, which is the probability of sharing zero IBD. The red dashed line represents the expected IBD. The black circles represent the samples in the NIDDK Non Jewish CD Controls and NIDDK GRU UC Cases. The results show that there are two pairs of related individuals in the merged data. A member from each pair was removed from the merged data.

iv. Genetic Analysis

The LDU mapping approach described in the Methods section of Chapter II was applied to the WTCCC2 Ulcerative Case-Control dataset, the merged NIDDK IBDO and GRU UC Cases and PD GWAS controls genotyped on the Illumina HumanHap550v.3 platform, and finally the merged NIDDK Non-Jewish CD controls and NIDDK GRU UC Cases genotyped on the Illumina HumanHap300v.1 and Illumina HumanHap300v.2 platforms respectively.

Each chromosome was divided into approximately 10 LDU non-overlapping windows for each dataset. Approximately the same LDU locations for the boundaries used for the Crohn's Disease analyses described in Chapters II and III were used for the UC datasets. This was done in order to facilitate the comparison of significant gene regions between the two subtypes IBD subtypes, which was to be our next aim. All the SNPs within a window were simultaneously tested within a composite likelihood (Λ) framework. The P-value was obtained, together with the estimated location $\hat{S}$ of the putative causal agent in LDU. The latter was then converted into kb by linear interpolation of the two flanking SNPs in HapMap.

v. Replications and meta-analysis of shared locations

The meta-analysis of the P-values was carried out if the analytical windows showed nominal statistical significance for at least 2 datasets and the estimated locations $\hat{S}$ for these nominally significant windows were within 80 kb of one another, which is a more stringent distance threshold for replication than was used for the CD GWAS described Chapter III since the UC data contain a greater SNP density.

The meta-analysis of the significant P-values obtained from the independent datasets was done using Fisher's combined probability test (see Methods section of Chapter III). The P-value for each test was combined as follows in order to produce a meta-analysis χ^2 with *2k* degrees of freedom for *k* number of tests:

$$\chi^2 = -2\sum_{i=1}^{k} ln(P)$$

A Bonferroni threshold was calculated as follows in order to account for multiple testing (see Methods section of Chapter II):

$$P - value\ threshold = \frac{\alpha}{Number\ of\ tests}$$

where α is the commonly accepted probability of 5% false positive rate. The Bonferroni P-value threshold used was 1×10^{-5}. This P-value threshold

corresponds to the total number of genome-wide analytical windows that were carried out for this analysis.

vi. Significant gene-regions identified

As described in Chapter III, the gene closest to the estimated location $\hat{S}$ was retrieved from a list of genome-wide RefSeq genes for each of the replicated analytical windows. If the distance between the gene retrieved and the estimated location $\hat{S}$ was larger than 300 kb, the estimated location $\hat{S}$ is listed as being "intergenic". Using BioMart, the Gene Ontology (GO) annotations attributed to each gene were retrieved and the results are shown in Appendix V.

III. RESULTS

i. Identifying the previously reported UC intervals.

As for the GWAS described in Chapter III of this thesis, the first step was to identify whether the LDU mapping method could detect the 59 previously reported loci. 47 of these 59 loci had been identified in a meta-analysis that consisted of six independent GWAS data and approximately 1.4 million autosomal markers. The remaining 12 loci were identified in the latest IBD meta-analysis, which consisted of fifteen GWAS data on CD and/or UC, as well as Immunochip data for approximately 75,000 cases and controls. In order for a previous locus to be considered as replicated using the LDU mapping method, the estimated location $\hat{S}$ or the 95% CI had to lie within, or overlap in the case of the 95% CI, the reported LD interval. This approach was taken since the genes represented in both published studies are the result of *in-silico* analyses that aimed to identify positional candidate genes of interest, rather than the certain causal location with direct evidence from functional studies. Out of the 59 previously identified loci, 53 were also found in at least one dataset within this study using an uncorrected significance threshold of P-value lower than 0.05. Six intervals showed no evidence of association with any of the three independent data sets. The statistical evidence and location estimates for the 53 loci are illustrated in Table 12. Most of the signals were identified in the larger WTCCC UC data set (49 out of 53 signals) and 65% of these (32 of the 49 signals)

survived the genome-wide significance Bonferroni threshold of $1x10^{-05.}$ 40% of the signals (21 out of 53) replicated in two datasets and 8% (4) replicated in all three datasets, with an estimated location $\hat{S}$ within 80 kb between all three datasets. Three of the 53 signals were replicated in the NIDDK IBDO + GRU dataset genotyped on the Illumina HumanHap 550v.3 (referred to as NIDDK 550 in Table 12) alone. In most cases, the previously reported intervals were quite large in size and the estimated locations $\hat{S}$ in the replication data sets were not within 80 kb of the other signal(s). In those cases it was thus not possible to meta-analyse the P-values for the individual datasets and the individual P-values for the datasets are shown in Table 12. Indeed, differences in estimated locations $\hat{S}$ larger than 80 kb could suggest the presence of more than one signal. This could be the case for the signal on Chromosome 11q (Table 12), shown in Figure 26. The previously reported interval spanned half a megabase in size and the LDU mapping approach further refined this interval with the identification of two signals using the WTCCC UC data within this interval and one using the NIDDK UC 550 data. The estimated locations for both datasets are located immediately upstream or within *MAML2* and in a region of LD breakdown (Figure 26). Considering the LD breakdown in the region, the dense SNP coverage in both datasets, and that none of the 95% CI overlap between the different datasets, it is unlikely that these signals represent one location and could be an indication for the presence of allelic heterogeneity. In such a case, meta-analysing P-values would imply the presence of only one signal and thus mask the presence of heterogeneity.

For 49% of the 53 loci identified (26 out of 53), the estimated location $\hat{S}$ is within a gene. The Consortium had reported positional candidate genes for 33 loci out of the 59 identified loci and 62% (16 out of 33 loci) of the genes illustrated in Table 12 match the genes reported by the Consortium for these intervals.

Since the reported LD intervals are quite large, in the instances where replication points to more than one gene it is difficult to determine whether this could be attributable to genuine genetic heterogeneity or simply an artifact of SNP resolution between the datasets. In Table 12 it can be seen that although some signals were identified using more than one dataset, the estimated location $\hat{S}$ for the datasets points to the same gene. 19 of the intervals shown in Table 12 have also been found to be shared with CD (Anderson et al., 2011) (see Table 12).

Table 12. Whole-genome association statistics and closest gene to the estimated location $\hat{S}$ of the causal variant for the previously reported UC loci.

Chr	Reported LD Interval (Mb)	WTCCC $\hat{S}$ (kb)	WTCCC P-value	NIDDK 550 $\hat{S}$ (kb)	NIDDK 550 P-value	NIDDK 300 $\hat{S}$ (kb)	NIDDK 300 P-value	WTCCC UC-NIDDK 550 $\hat{S}$ difference (kb)	WTCCC UC-NIDDK 300 $\hat{S}$ difference (kb)	NIDDK 550-NIDDK 300 $\hat{S}$ difference (kb)	Meta P-value	Gene
1p	2.40-2.81	2,473.7	$2.1x10^{-05}$	-	-	2,469.2	$2.3x10^{-02}$	-	4.5	-	$7.5x10^{-06}$	*HES5/LOC115110*
1p	7.91-8.21	-	-	8,270.0	$1.8x10^{-02}$	7,921.9	$5.6x10^{-03}$	-	-	348.1	-	*UTS2/SLC45A1*
1p	20.06-20.31	20,213.3	$7.3x10^{-11}$	20,169.9	$7.6x10^{-06}$	20,141.4	$1.7x10^{-03}$	43.5	72.0	28.5	$8.3x10^{-16}$	*OTUD3*/RNF186**
1p	22.67-22.74	22,695.0	$2.3x10^{-02}$	-	-	-	-	-	-	-	-	*ZBTB40*
1p	67.53-67.77	67,769.1	$1.2x10^{-04}$	-	-	67,741.5	$3.5x10^{-05}$	-	27.6	-	$8.4x10^{-08}$	*IL23R*§*
1q	161.27-161.64	161,409.4	$1.4x10^{-03}$	-	-	161,479.8	$6.9x10^{-03}$	-	70.4	-	$1.2x10^{-04}$	*FCGR2A**
1q	200.79-201.06	201,015.5	$2.5x10^{-19}$	201,041.9	$4.7x10^{-03}$	201,012.1	$6.5x10^{-04}$	26.3	3.5	29.8	$1.2x10^{-21}$	*CACNA1S*§*
1q	206.78-207.04	206,932.3	$2.9x10^{-05}$	-	-	206,939.9	$2.1x10^{-03}$	-	7.6	-	$1.1x10^{-06}$	*IL10§*
2p	60.91-62.02	61,186.8	$1.3x10^{-09}$	-	-	61,181.8	$7.0x10^{-03}$	-	5.0	-	$2.3x10^{-10}$	*PUS10*§*
2q	102.29-102.76	102,170.4	$5.1x10^{-05}$	-	-	102,733.0	$3.7x10^{-03}$	-	562.7	-	-	*IL1R1/RFX8§*
2q	199.27-200.10	199,497.6	$4.9x10^{-17}$	-	-	-	-	-	-	-	-	*PLCL1*
2q	218.87-219.26	219,001.3	$7.0x10^{-06}$	218,928.7	$4.7x10^{-04}$	218,972.1	$1.5x10^{-08}$	72.7	29.3	43.4	$3.7x10^{-14}$	*CXCR2*/RUFY4**
2q	241.55-241.67	241,563.4	$1.8x10^{-06}$	-	-	-	-	-	-	-	-	*GPR35**
3p	48.16-51.79	49,894.9	$3.4x10^{-53}$	48,687.8	$4.1x10^{-09}$	-	-	1,207.1	-	-	-	*CELSR3*/CAMKV§*
3p	52.80-53.30	52,833.5	$3.2x10^{-13}$	53,252.5	$2.2x10^{-02}$	-	-	419.1	-	-	-	*ITIH3*/TKT*
4q	103.26-103.76	103,419.3	$2.2x10^{-12}$	-	-	104,361.3	$5.5x10^{-08}$	-	942.5	-	-	*NFKB1/TACR3*
4q	122.98-123.56	123,025.1	$6.5x10^{-09}$	-	-	-	-	-	-	-	-	*KIAA1109*
5p	0.43-0.75	562.8	$9.2x10^{-09}$	578.8	$2.0x10^{-03}$	-	-	15.9	-	-	$4.8x10^{-10}$	*SLC9A3/CEP72*
5p	10.67-10.85	-	-	10,710.8	$3.0x10^{-06}$	-	-	-	-	-	-	*DAP**

Chr	Reported LD Interval (Mb)	WTCCC $\hat{S}$ (kb)	WTCCC P-value	NIDDK 550 $\hat{S}$ (kb)	NIDDK 550 P-value	NIDDK 300 $\hat{S}$ (kb)	NIDDK 300 P-value	WTCCC UC-NIDDK 550 $\hat{S}$ difference (kb)	WTCCC UC-NIDDK 300 $\hat{S}$ difference (kb)	NIDDK 550-NIDDK 300 $\hat{S}$ difference (kb)	Meta P-value	Gene
5p	35.79-36.03	35,811.4	2.5×10^{-12}	-	-	-	-	-	-	-	-	*SPEF2**
5p	40.28-40.81	-	-	40,617.6	3.3×10^{-13}	-	-	-	-	-	-	*PTGER4*
5q	134.38-134.50	134,463.1	7.2×10^{-05}	-	-	-	-	-	-	-	-	*C5orf66**
5q	158.53-158.93	158,825.8	1.1×10^{-02}	-	-	-	-	-	-	-	-	*IL12B*
6p	31.38-32.90	32,804.8	2.9×10^{-12}	32,739.0	3.0×10^{-14}	-	-	65.8	-	-	5.0×10^{-24}	*TAP2*/HLA-DQB2*
6q	137.84-138.13	137,978.5	4.0×10^{-08}	-	-	-	-	-	-	-	-	*OLIG3*
7p	2.73-2.93	2,759.0	3.0×10^{-07}	-	-	-	-	-	-	-	-	*AMZ1*
7p	26.97-27.47	27,412.2	6.4×10^{-08}	-	-	-	-	-	-	-	-	*EVX1*
7q	107.41-107.60	107,459.0	6.1×10^{-10}	-	-	-	-	-	-	-	-	*SLC26A3*
	-	107,476.9	9.9×10^{-27}	-	-	107,603.9	2.1×10^{-02}	-	127.0	-	-	*LAMB1*/SLC26A3*
7q	128.54-128.77	129,310.2	2.5×10^{-05}	-	-	-	-	-	-	-	-	*KCP*
9p	4.94-5.29	5,012.4	4.8×10^{-27}	-	-	4,911.4	1.1×10^{-02}	-	101.0	-	-	*RCL1/JAK2**
9q	117.44-117.70	117,559.8	2.2×10^{-13}	-	-	-	-	-	-	-	-	*TNFSF15**
9q	139.13-139.43	139,262.0	8.7×10^{-13}	-	-	-	-	-	-	-	-	*CARD9**
10p	35.18-35.90	35,554.4	8.8×10^{-05}	-	-	-	-	-	-	-	-	*CCNY**
10q	101.26-101.34	101,754.7	3.2×10^{-16}	101,342.5	3.1×10^{-02}	-	-	72.6	-	-	-	*NKX2-3/SLC25A28*
11q	76.04-76.34	76,303.3	3.5×10^{-07}	-	-	-	-	-	-	-	-	*C11orf30*
11q	95.77-96.27	95,762.1	3.1×10^{-02}	95,703.9	1.0×10^{-02}	-	-	58.2	-	-	2.9×10^{-03}	*MAML2**
	-	96,008.6	4.2×10^{-08}	-	-	-	-	-	-	-	-	*MAML2**
11q	114.25-114.57	114,331.5	4.3×10^{-07}	-	-	114,239.1	9.3×10^{-03}	-	92.4	-	-	*C11orf71/REXO2*
12q	68.43-68.63	68,478.2	6.7×10^{-11}	68,493.0	1.4×10^{-04}	68,416.4	1.1×10^{-02}	14.8	61.8	76.5	7.2×10^{-14}	*IFNG*
13q	27.49-27.56	27,543.0	3.3×10^{-17}	27,564.7	4.0×10^{-03}	-	-	21.6	-	-	6.0×10^{-18}	*USP12*
13q	41.00-41.05	41,025.9	1.4×10^{-03}	-	-	41,014.0	2.7×10^{-03}	-	11.9	-	4.9×10^{-05}	*LINC00598**
15q	41.29-41.81	41,799.3	7.0×10^{-15}	-	-	-	-	-	-	-	-	*LTK**

Chr	Reported LD Interval (Mb)	WTCCC $\hat{S}$ (kb)	WTCCC P-value	NIDDK 550 $\hat{S}$ (kb)	NIDDK 550 P-value	NIDDK 300 $\hat{S}$ (kb)	NIDDK 300 P-value	WTCCC UC-NIDDK 550 $\hat{S}$ difference (kb)	WTCCC UC-NIDDK 300 $\hat{S}$ difference (kb)	NIDDK 550-NIDDK 300 $\hat{S}$ difference (kb)	Meta P-value	Gene
16q	68.42-68.84	68,592.0	$8.9x10^{-13}$	-	-	-	-	-	-	-	-	*ZFP90*
16q	85.97-86.02	85,996.3	$1.5x10^{-04}$	-	-	-	-	-	-	-	-	*IRF8*
17q	37.37-38.26	37,756.0	$1.0x10^{-19}$	38,257.1	$1.3x10^{-02}$	-	-	501.1	-	-	-	*NR1D1/NEUROD2*
17q	70.39-70.89	70,697.0	$1.1x10^{-02}$	-	-	-	-	-	-	-	-	*SLC39A11**
19q	46.87-47.37	33,806.8	-	46,837.5	$5.2x10^{-03}$	-	-	-	-	-	-	*HIF3A**
20q	33.55-34.05	43,065.0	$2.5x10^{-09}$	-	-	-	-	-	-	-	-	*MMP24*
20q	43.06-43.27	-	$4.6x10^{-09}$	-	-	-	-	-	-	-	-	*HNF4A*
20q	62.19-62.51	62,326.9	$5.9x10^{-04}$	-	-	-	-	-	-	-	-	*RTEL1**
21q	16.70-16.85	16,818.4	$4.9x10^{-07}$	16,802.9	$5.6x10^{-05}$	-	-	15.5	-	-	$6.9x10^{-10}$	*USP25*
21q	40.42-40.49	40,474.0	$3.6x10^{-03}$	-	-	-	-	-	-	-	-	*PSMG1*
21q	45.59-45.70	45,608.8	$2.2x10^{-04}$	-	-	-	-	-	-	-	-	*ICOSLG*
22q	50.31-50.49	50,351.1	$3.4x10^{-02}$	-	-	-	-	-	-	-	-	*PIM3*

* denotes that at least one of the estimated locations $\hat{S}$ is located within the identified gene. The intervals highlighted in blue correspond to the loci previously identified in the IBD meta-analysis that also used the Immunochip genotyping array (Jostins et al., 2012). The intervals highlighted in grey correspond to loci that have been found to be shared with CD (Anderson et al., 2011). [§] The Meta P-value is only provided in instances where the $\hat{S}$ difference is less than 80 kb between datasets.

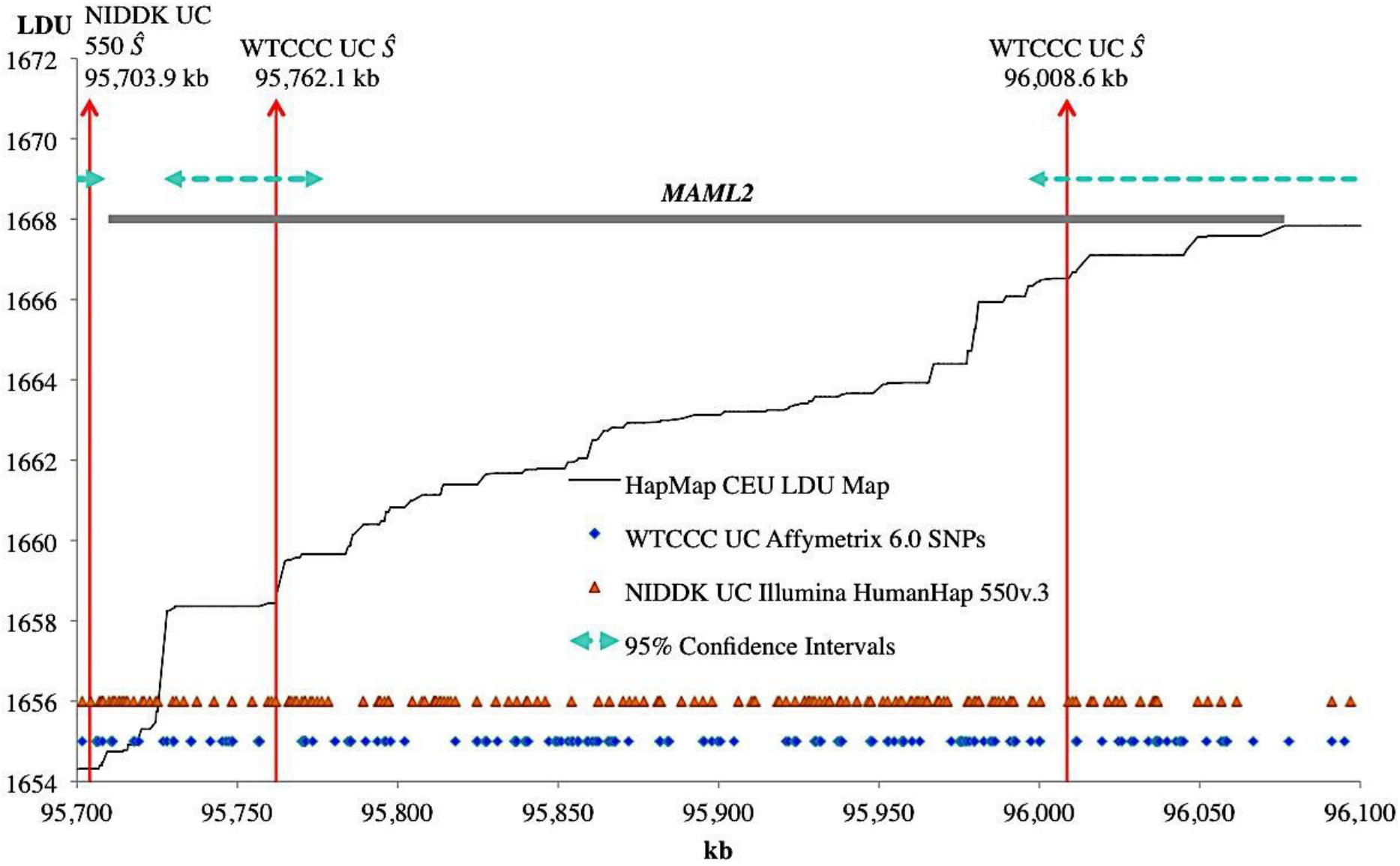

LDU
NIDDK UC
550 Ŝ
95,703.9 kb
WTCCC UC Ŝ
95,762.1 kb
WTCCC UC Ŝ
96,008.6 kb
MAML2
1672
1670
1668
1666
1664
1662
1660
1658
1656
1654
HapMap CEU LDU Map
WTCCC UC Affymetrix 6.0 SNPs
NIDDK UC Illumina HumanHap 550v.3
95% Confidence Intervals
95,700
95,750
95,800
95,850
95,900
95,950
96,000
96,050
96,100
kb

Figure 26. Localisation within the *MAML2* region.
The Y-axis represents genomic LDU distances and the X-axis represents physical distances in kilobases. The solid black line represents the LDU map of the region constructed from the HapMap CEU data. The red vertical arrow is the estimated location $\hat{S}$ of the causal variants. The blue diamonds show the SNPs found in this regions on the WTCCC UC Affymetrix 6.0 array. The light green horizontal dashed lines represent the 95% CIs of the estimated locations $\hat{S}$. The orange triangles are the SNPs on the NIDDK UC Illumina HumanHap 550v.3 array in this region.

i. 138 novel gene-regions identified

The second step was to determine whether the LDU mapping method would be able to detect novel signals using the three UC case-controls datasets. For this a meta-analysis approach was used in which the maximum distance between the estimated $\hat{S}$ locations was set at 80kb. 138 novel gene-regions that passed the Bonferroni genome-wide significance threshold of $1x10^{-05}$ were identified. Since the distance between the estimated locations $\hat{S}$ was less than 80 kb, the meta-analysis P-value is provided for these replications. 10 of these 138 gene-regions replicate with all three datasets, with a meta P-value surviving the Bonferroni genome-wide significance threshold. One of these ten replications, the *ERAP1* gene-region, is illustrated in Figure 27. For this region, the estimated $\hat{S}$ locations for the WTCCC UC and NIDDK UC Illumina HumanHap 550v.3 datasets lie within intron 17 of the *ERAP1* gene whereas, although still very close, the estimated $\hat{S}$ location of the NIDDK UC Illumina HumanHap 300v.1 & v.2 dataset lies downstream of the gene. The estimated locations $\hat{S}$ and their 95% CI are illustrated in Figure 27.

The remaining 128 novel gene-regions were replicated in two out of three datasets, with a meta P-value surviving the Bonferroni genome-wide significance threshold.

Out of these 138 gene regions, 12 signals were intergenic locations, where the closest gene was further than 300kb from the estimated locations $\hat{S}$. Of the

remaining 126 gene-regions, 57% of the signals were intragenic locations, with the estimated location $\hat{S}$ being within a gene for at least one of the datasets. Additionally, in 75% of the 126 gene-regions, the estimated locations $\hat{S}$ for the datasets point to the same gene.

Table 13. Whole-genome association statistics and closest gene to the estimated location $\hat{S}$ of the causal variant for the 10 novel gene-regions that replicate in all three datasets with estimated locations $\hat{S}$ within 80 kb of each other.

Chr	WTCCC UC	NIDDK IBDO + GRU Illumina 550v.3	NIDDK GRU UC Illumina 300	Meta P-value	Gene
1q	171,661.9	171,697.8	171,620.2	$8.3x10^{-09}$	*VAMP4*/MYOC**
3q	109,508.6	109,515.6	109,521.4	$2.0x10^{-11}$	*intergenic*
4p	30,242.7	30,306.9	30,277.9	$9.3x10^{-07}$	*intergenic*
4q	101,853.1	101,798.3	101,799.1	$1.0x10^{-15}$	*PPP3CA*
5q	96,109.8	96,108.9	96,134.9	$8.9x10^{-07}$	*ERAP1**
6p	33,051.2	32,998.7	33,051.2	$2.0x10^{-10}$	*HLA-DPB1*/HLA-DOA*
7q	144,299.5	144,249.7	144,297.3	$3.3x10^{-07}$	*TPK1**
10p	14,105.9	14,066.5	14,108.0	$6.7x10^{-09}$	*FRMD4A**
12p	28,679.5	28,595.9	28,309.1	$7.3x10^{-27}$	***CCDC91****
20q	31,663.2	31,722.5	31,682.8	$2.1x10^{-09}$	*BPIFB3/BPIFB4**

* denotes that at least one of the estimated locations $\hat{S}$ is located within the identified gene. Genes in bold are genes that have been identified in Chapter III for CD. The gene region highlighted in grey is included in this table even though the estimated locations are further than 80 kb from one another because all the estimated locations for this gene-region are genetically very close as measure in LDU, due to its occurrence within a long region of very high LD).

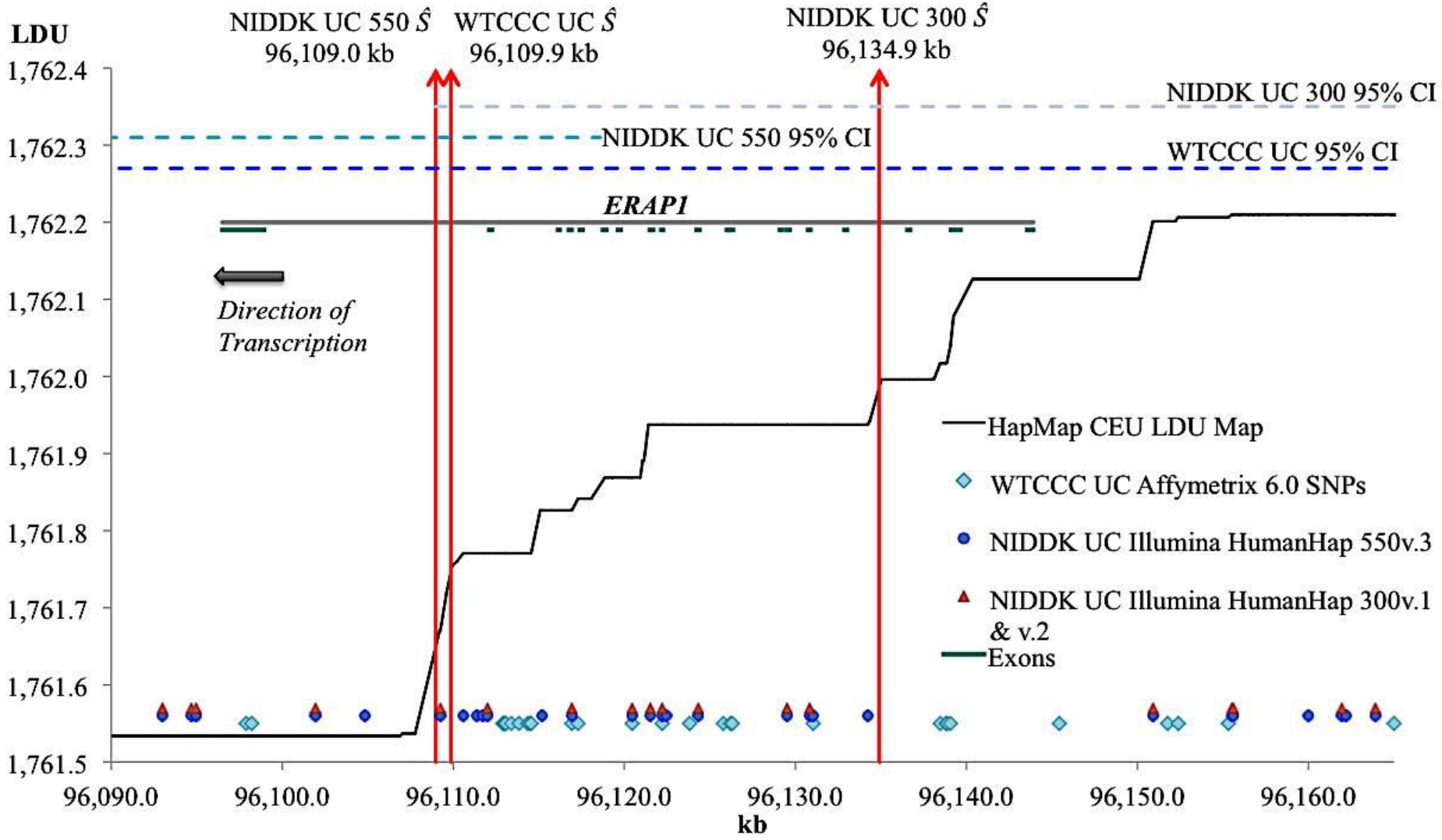

LDU
NIDDK UC 550 Ŝ
96,109.0 kb
WTCCC UC Ŝ
96,109.9 kb
NIDDK UC 300 Ŝ
96,134.9 kb
NIDDK UC 300 95% CI
NIDDK UC 550 95% CI
WTCCC UC 95% CI
ERAP1
Direction of Transcription
1,762.4
1,762.3
1,762.2
1,762.1
1,762.0
1,761.9
1,761.8
1,761.7
1,761.6
1,761.5
96,090.0
96,100.0
96,110.0
96,120.0
96,130.0
96,140.0
96,150.0
96,160.0
kb
HapMap CEU LDU Map
WTCCC UC Affymetrix 6.0 SNPs
NIDDK UC Illumina HumanHap 550v.3
NIDDK UC Illumina HumanHap 300v.1 & v.2
Exons

Figure 27. Localisation within the *ERAP1* region.
The Y-axis represents genomic LDU locations whereas the X-axis represents genomic physical locations in kilobases. The solid black line represents the LDU map of the region constructed from the HapMap CEU data. The red vertical arrows are the estimated locations $\hat{S}$ of the causal variants. The dark green lines highlight the exons of the *ERAP1* gene. The dark blue bashed line reporesentes the 95% CI for the WTCCC UC estimated location $\hat{S}$. The light blue dashed line represents the 95% CI for the NIDDK UC Illumina HumanHap 550v.3 estimated location $\hat{S}$. The grey dashed line represents the 95% CI of the NIDDK UC Illumina HumanHap 300v.1&v.2 estimated location $\hat{S}$. The light blue diamonds show the SNPs found in this regions on the WTCCC UC Affymetrix 6.0 array, the dark blue circles represent those on the NIDDK UC Illumina HumanHap 550v.3 array and the red triangles represent the SNPs in this regions that are common to both the NIDDK Illumina HumanHap 300v.1 and Illumina HumanHap 300v.2 arrays.

Table 14. Whole-genome association statistics and closest gene to the estimated location $\hat{S}$ of the causal variant for the 128 novel gene-regions that replicate with two datasets with an estimated location $\hat{S}$ within 80 kb between datasets. a) Whole-genome association statistics and closest gene to the estiamted location $\hat{S}$ of the causal variant for the 108 novel gene-regions that replicate with two datasets with an estimated location $\hat{S}$ within 80 kb for WTCCC UC and NIDDK UC 550 or WTCCC UC and NIDDK UC 300 datasets. b) Whole-genome association statistics and closest gene to the estiamted location $\hat{S}$ of the causal variant for the 20 novel gene-regions that replicate with two datasets with an estimated location $\hat{S}$ within 80 kb for NIDDK UC 550 and NIDDK UC 300 datasets.

Table 14 a)

Chr	UC Replication (Signal 1/Signal 2)	Signal 1	Signal 2	Meta P-value	Gene
1p	WTCCC/NIDDK 300	26,531.4	26,528.4	$3.4x10^{-10}$	*CATSPER4**
1p	WTCCC/NIDDK 550	79,063.7	78,988.3	$3.3x10^{-12}$	*IFI44L/PTGFR**
1p	WTCCC/NIDDK 550	92,316.2	92,316.9	$1.0x10^{-08}$	*TGFBR3**
1p	WTCCC/NIDDK 300	101,501.7	101,500.7	$4.0x10^{-11}$	*DPH5*
1q	WTCCC/NIDDK 300	151,907.1	151,985.6	$8.3x10^{-17}$	*THEM4/NBPF18P*
1q	WTCCC/NIDDK 300	173,304.4	173,370.2	$4.6x10^{-07}$	*LOC100506023**
1q	WTCCC/NIDDK 550	232,574.7	232,566.9	$7.6x10^{-07}$	*SIPA1L2**
1q	WTCCC/NIDDK 550	236,586.9	236,539.8	$1.0x10^{-05}$	*EDARADD**
2p	WTCCC/NIDDK 550	3,006.2	2,935.2	$3.2x10^{-06}$	*TSSC1*
2p	WTCCC/NIDDK 550	15,189.0	15,112.8	$3.3x10^{-07}$	*NBAS*
2p	WTCCC/NIDDK 550	44,039.9	44,002.0	$7.3x10^{-08}$	*ABCG5*/DYNC2LI1**
2p	WTCCC/NIDDK 300	49,056.4	49,085.9	$7.2x10^{-08}$	*STON1-GTF2A1L*
2p	WTCCC/NIDDK 300	80,620.5	80,665.6	$7.7x10^{-08}$	*CTNNA2**
2p	WTCCC/NIDDK 300	81,633.5	81,665.4	$2.3x10^{-08}$	*intergenic*
2q	WTCCC/NIDDK 550	100,805.3	100,871.1	$1.1x10^{-10}$	*LOC150577*
2q	WTCCC/NIDDK 300	103,125.4	103,159.7	$5.7x10^{-26}$	*SLC9A4**
2q	WTCCC/NIDDK 550	160,866.9	160,889.9	$1.9x10^{-07}$	*PLA2R1**
2q	WTCCC/NIDDK 550	168,748.0	168,738.6	$7.5x10^{-06}$	*B3GALT1*
2q	WTCCC/NIDDK 550	169,969.0	170,001.5	$4.7x10^{-06}$	***LRP2****
2q	WTCCC/NIDDK 550	226,257.7	226,231.6	$8.8x10^{-08}$	***NYAP2***
2q	WTCCC/NIDDK 550	237,052.2	237,064.0	$3.3x10^{-06}$	*AGAP1/GBX2*
3p	WTCCC/NIDDK 550	29,707.3	29,707.6	$1.1x10^{-09}$	*RBMS3**
3p	WTCCC/NIDDK 300	30,626.6	30,680.0	$2.7x10^{-07}$	*TGFBR2**
3p	WTCCC/NIDDK 550	60,514.8	60,478.9	$3.8x10^{-06}$	*FHIT**
3p	WTCCC/NIDDK 300	61,299.4	61,267.2	$9.0x10^{-08}$	*FHIT*
3q	WTCCC/NIDDK 550	156,277.0	156,274.8	$9.7x10^{-06}$	*SSR3*

Chr	UC Replication (Signal 1/Signal 2)	Signal 1	Signal 2	Meta P-value	Gene
3q	WTCCC/NIDDK 550	197,297.5	197,240.1	$1.1x10^{-07}$	*BDH1**
4p	WTCCC/NIDDK 550	36,825.2	36,753.0	$2.7x10^{-06}$	*intergenic*
4q	WTCCC/NIDDK 300	84,227.2	84,162.2	$6.1x10^{-07}$	*HPSE*/COQ2*
4q	WTCCC/NIDDK 550	124,451.7	124,448.6	$2.7x10^{-08}$	*LOC285419/SPRY1*
4q	WTCCC/NIDDK 300	163,012.0	163,030.3	$1.4x10^{-06}$	*FSTL5**
5p	WTCCC/NIDDK 550	2,740.1	2,774.8	$5.0x10^{-06}$	*IRX2/C5orf38*
5p	WTCCC/NIDDK 550	3,176.2	3,114.8	$8.5x10^{-07}$	*LINC01019*
5p	WTCCC/NIDDK 300	9,782.2	9,824.2	$1.5x10^{-08}$	*LOC285692**
5p	WTCCC/NIDDK 550	15,752.0	15,789.8	$1.6x10^{-07}$	*FBXL7**
5p	WTCCC/NIDDK 550	28,770.4	28,718.5	$2.6x10^{-07}$	*LSP1P3*
5p	WTCCC/NIDDK 550	39,115.8	39,195.1	$8.4x10^{-09}$	*FYB**
5q	WTCCC/NIDDK 550	115,387.3	115,419.8	$5.6x10^{-08}$	*ARL14EPL*/COMMD10*
5q	WTCCC/NIDDK 550	132,476.0	132,529.7	$8.3x10^{-07}$	*HSPA4/FSTL4*
5q	WTCCC/NIDDK 300	139,269.1	139,264.8	$4.3x10^{-06}$	*NRG2**
5q	WTCCC/NIDDK 550	143,483.6	143,484.5	$9.9x10^{-07}$	*YIPF5*
5q	WTCCC/NIDDK 550	179,756.8	179,816.5	$6.4x10^{-06}$	*GFPT2**
6p	WTCCC/NIDDK 550	12,353.2	12,297.1	$5.1x10^{-08}$	*EDN1**
6p	WTCCC/NIDDK 300	12,951.2	13,029.5	$4.3x10^{-06}$	*PHACTR1**
6q	WTCCC/NIDDK 550	70,361.8	70,362.8	$2.3x10^{-11}$	*LMBRD1*
6q	WTCCC/NIDDK 550	86,817.3	86,798.5	$4.6x10^{-23}$	*MIR548AD*
6q	WTCCC/NIDDK 550	103,900.6	103,905.8	$5.2x10^{-09}$	*HACE1*
6q	WTCCC/NIDDK 550	119,030.3	119,029.7	$2.8x10^{-15}$	*CEP85L**
7p	WTCCC/NIDDK 300	16,695.7	16,710.5	$5.9x10^{-07}$	*BZW2**
7p	WTCCC/NIDDK 300	41,923.9	41,941.0	$1.4x10^{-10}$	***GLI3***
7q	WTCCC/NIDDK 550	67,238.6	67,292.7	$6.7x10^{-07}$	*intergenic*
7q	WTCCC/NIDDK 300	80,060.8	80,056.9	$3.7x10^{-06}$	*GNAT3*
7q	WTCCC/NIDDK 550	102,791.0	102,791.4	$4.4x10^{-10}$	*NAPEPLD*
7q	WTCCC/NIDDK 550	108,774.1	108,747.1	$3.7x10^{-10}$	*C7orf66*
7q	WTCCC/NIDDK 300	146,952.3	146,936.0	$3.3x10^{-10}$	*CNTNAP2**
7q	WTCCC/NIDDK 550	147,981.9	147,927.4	$1.8x10^{-06}$	*MIR548T**
8p	WTCCC/NIDDK 550	6,758.4	6,789.3	$2.0x10^{-08}$	*DEFB1/DEFA4*
8p	WTCCC/NIDDK 550	24,858.2	24,844.8	$1.0x10^{-08}$	*NEFL*
8q	WTCCC/NIDDK 550	56,505.8	56,523.5	$1.9x10^{-09}$	*XKR4*
8q	WTCCC/NIDDK 300	60,327.6	60,254.8	$8.1x10^{-06}$	*TOX*
9p	WTCCC/NIDDK 550	1,317.7	1,237.9	$2.4x10^{-07}$	*DMRT2*
9p	WTCCC/NIDDK 300	14,404.6	14,446.0	$6.2x10^{-06}$	*NFIB*
9p	WTCCC/NIDDK 550	30,225.6	30,292.2	$1.1x10^{-08}$	*LOC401497*
9q	WTCCC/NIDDK 300	81,509.0	81,572.2	$1.5x10^{-06}$	*intergenic*
9q	WTCCC/NIDDK 300	90,393.1	90,420.4	$1.7x10^{-08}$	*CTSL3P**
9q	WTCCC/NIDDK 300	103,473.9	103,545.4	$3.0x10^{-08}$	*MURC*

Chr	UC Replication (Signal 1/Signal 2)	Signal 1	Signal 2	Meta P-value	Gene
9q	WTCCC/NIDDK 550	115,736.2	115,723.4	$2.6x10^{-06}$	*ZNF883*
9q	WTCCC/NIDDK 550	120,967.8	121,009.7	$1.2x10^{-07}$	*intergenic*
10p	WTCCC/NIDDK 300	36,971.1	36,941.3	$1.2x10^{-08}$	*ANKRD30A*
10q	WTCCC/NIDDK 550	63,552.9	63,553.5	$1.7x10^{-08}$	***C10orf107***
10q	WTCCC/NIDDK 550	90,739.8	90,787.4	$6.0x10^{-07}$	*ACTA2*/**FAS***
10q	WTCCC/NIDDK 550	102,848.6	102,880.7	$7.5x10^{-06}$	*TLX1NB**
10q	WTCCC/NIDDK 550	104,232.3	104,210.8	$3.3x10^{-14}$	*TMEM180*/C10orf95**
10q	WTCCC/NIDDK 550	121,292.5	121,287.8	$2.6x10^{-06}$	*RGS10**
11p	WTCCC/NIDDK 300	13,317.5	13,353.6	$6.7x10^{-10}$	*ARNTL**
11p	WTCCC/NIDDK 300	15,470.0	15,402.1	$8.8x10^{-09}$	*INSC*
11p	WTCCC/NIDDK 550	17,426.6	17,426.3	$8.1x10^{-13}$	*ABCC8**
11p	WTCCC/NIDDK 550	20,160.1	20,124.2	$7.7x10^{-08}$	*NAV2**
11p	WTCCC/NIDDK 300	42,635.2	42,657.9	$3.7x10^{-07}$	*intergenic*
11q	WTCCC/NIDDK 550	77,904.5	77,918.6	$3.4x10^{-21}$	*USP35**
11q	WTCCC/NIDDK 550	102,575.7	102,575.7	$1.3x10^{-12}$	*MMP27**
11q	WTCCC/NIDDK 550	117,040.9	116,994.7	$3.9x10^{-09}$	*PAFAH1B2**
12p	WTCCC/NIDDK 550	32,931.7	32,998.7	$2.6x10^{-07}$	*PKP2/PKP2**
12q	WTCCC/NIDDK 300	73,009.6	72,984.5	$2.5x10^{-07}$	*TRHDE**
12q	WTCCC/NIDDK 550	108,637.3	108,693.5	$2.6x10^{-07}$	*WSCD2*/CMKLR1**
12q	WTCCC/NIDDK 550	126,545.5	126,511.6	$7.6x10^{-07}$	*LINC00939*
13q	WTCCC/NIDDK 550	40,510.3	40,546.3	$2.2x10^{-12}$	*COG6*
13q	WTCCC/NIDDK 300	71,015.4	71,012.4	$1.5x10^{-07}$	*ATXN8OS*
14q	WTCCC/NIDDK 300	27,211.4	27,271.3	3.4x10-18	*NOVA1/MIR4307*
14q	WTCCC/NIDDK 550	40,736.6	40,766.6	$6.1x10^{-10}$	*intergenic*
14q	WTCCC/NIDDK 550	56,435.8	56,498.9	$3.2x10^{-06}$	*PELI2*
14q	WTCCC/NIDDK 550	62,955.4	62,954.2	$5.6x10^{-09}$	*KCNH5*
14q	WTCCC/NIDDK 550	85,994.1	86,053.4	$9.4x10^{-07}$	*FLRT2**
14q	WTCCC/NIDDK 550	96,023.1	96,032.9	$1.4x10^{-06}$	*GLRX5*
15q	WTCCC/NIDDK 300	47,995.0	47,981.0	$1.0x10^{-06}$	*SEMA6D**
15q	WTCCC/NIDDK 550	61,432.1	61,378.3	$2.0x10^{-07}$	*RORA**
15q	WTCCC/NIDDK 550	63,050.8	63,098.4	$2.1x10^{-09}$	*TLN2**
15q	WTCCC/NIDDK 550	63,819.8	63,758.9	$2.0x10^{-10}$	*USP3**
16p	WTCCC/NIDDK 550	20,025.2	20,036.1	$6.8x10^{-08}$	*GPR139*
17q	WTCCC/NIDDK 550	70,080.6	70,145.4	$2.7x10^{-09}$	*FLJ37644*/SOX9*
18q	WTCCC/NIDDK 300	30,889.1	30,891.0	$9.9x10^{-11}$	*CCDC178**
18q	WTCCC/NIDDK 550	39,026.9	39,010.8	$8.8x10^{-06}$	*KC6*
18q	WTCCC/NIDDK 550	46,156.0	46,112.6	$8.4x10^{-08}$	*CTIF**
18q	WTCCC/NIDDK 300	67,510.4	67,442.5	$5.3x10^{-06}$	*DOK6**
20p	WTCCC/NIDDK 550	14,314.2	14,352.4	$1.2x10^{-17}$	*FLRT3*/**MACROD2***
20q	WTCCC/NIDDK 550	45,577.9	45,571.8	$1.9x10^{-08}$	*EYA2**

Chr	UC Replication (Signal 1/Signal 2)	Signal 1	Signal 2	Meta P-value	Gene
21q	WTCCC/NIDDK 550	26,833.1	26,770.9	$3.5x10^{-17}$	*LINC00158**
22q	WTCCC/NIDDK 550	34,860.0	34,784.2	$9.7x10^{-07}$	*intergenic*

Table 14 b)

Chr	UC Replication (Signal 1/Signal 2)	Signal 1	Signal 2	Meta P-value	Gene
1q	NIDDK 550/NIDDK 300	177,964.6	178,030.4	$8.7x10^{-06}$	*LOC730102*
1q	NIDDK 550/NIDDK 300	226,845.3	226,846.9	$5.6x10^{-06}$	*ITPKB-IT1**
2p	NIDDK 550/NIDDK 300	29,978.8	29,979.3	$2.6x10^{-07}$	*ALK**
2q	NIDDK 550/NIDDK 300	134,975.7	135,031.4	$4.1x10^{-10}$	*MGAT5**
2q	NIDDK 550/NIDDK 300	137,187.7	137,187.0	$8.0x10^{-27}$	*CXCR4*
2q	NIDDK 550/NIDDK 300	181,752.0	181,676.6	$1.2x10^{-07}$	*UBE2E3*
3q	NIDDK 550/NIDDK 300	114,511.3	114,511.1	$1.5x10^{-07}$	*ZBTB20**
4p	NIDDK 550/NIDDK 300	38,734.2	38,707.6	$1.6x10^{-19}$	*KLF3*
4q	NIDDK 550/NIDDK 300	148,723.6	148,722.9	$3.6x10^{-07}$	*ARHGAP10**
5p	NIDDK 550/NIDDK 300	30,570.2	30,547.0	$3.3x10^{-07}$	*intergenic*
5q	NIDDK 550/NIDDK 300	170,170.2	170,212.3	$3.5x10^{-06}$	*KCNIP1/GABRP**
5q	NIDDK 550/NIDDK 300	176,784.7	176,827.1	$2.2x10^{-07}$	*RGS14/PFN3**
10q	NIDDK 550/NIDDK 300	52,198.4	52,261.6	$1.4x10^{-08}$	*SGMS1**
12p	NIDDK 550/NIDDK 300	1,495.7	1,509.9	$1.3x10^{-06}$	*ERC1**
12q	NIDDK 550/NIDDK 300	119,858.9	119,858.9	$3.8x10^{-06}$	*CCDC60**
15q	NIDDK 550/NIDDK 300	72,677.4	72,710.7	$3.4x10^{-12}$	*HEXA-AS1/TMEM202*
16q	NIDDK 550/NIDDK 300	82,279.3	82,278.8	$4.6x10^{-08}$	*MPHOSPH6*
18q	NIDDK 550/NIDDK 300	41,809.4	41,810.7	$4.8x10^{-08}$	*intergenic*
18q	NIDDK 550/NIDDK 300	59,990.6	60,005.0	$2.8x10^{-09}$	*TNFRSF11A**
20q	NIDDK 550/NIDDK 300	54,350.0	54,421.0	$8.3x10^{-06}$	*CBLN4*

* denotes that at least one of the estimated locations $\hat{S}$ is located within the identified gene. Genes in bold are genes that have been identified in Chapter III for CD.

From the 128 novel gene-regions illustrated in Table 14 a) and b), 68% of the signals have an estimated location $\hat{S}$ less than 50 kb apart in the two datasets and 40% had an estimated location $\hat{S}$ within less than 25 kb of each other. An example is shown in Figure 28, which illustrates the *MMP27* gene-region. The estimated locations $\hat{S}$ for the WTCCC UC data and the NIDDK UC Illumina HumanHap 500v.3 data are identical for this region and are located in a small block of LD. Additionally, 7 of the 138 novel gene-regions were also identified in the CD analysis carried out in Chapter III, as shown in bold in Table 13 and 13. The molecular and functional attributes of all the genes identified in this chapter are shown in Appendix V.

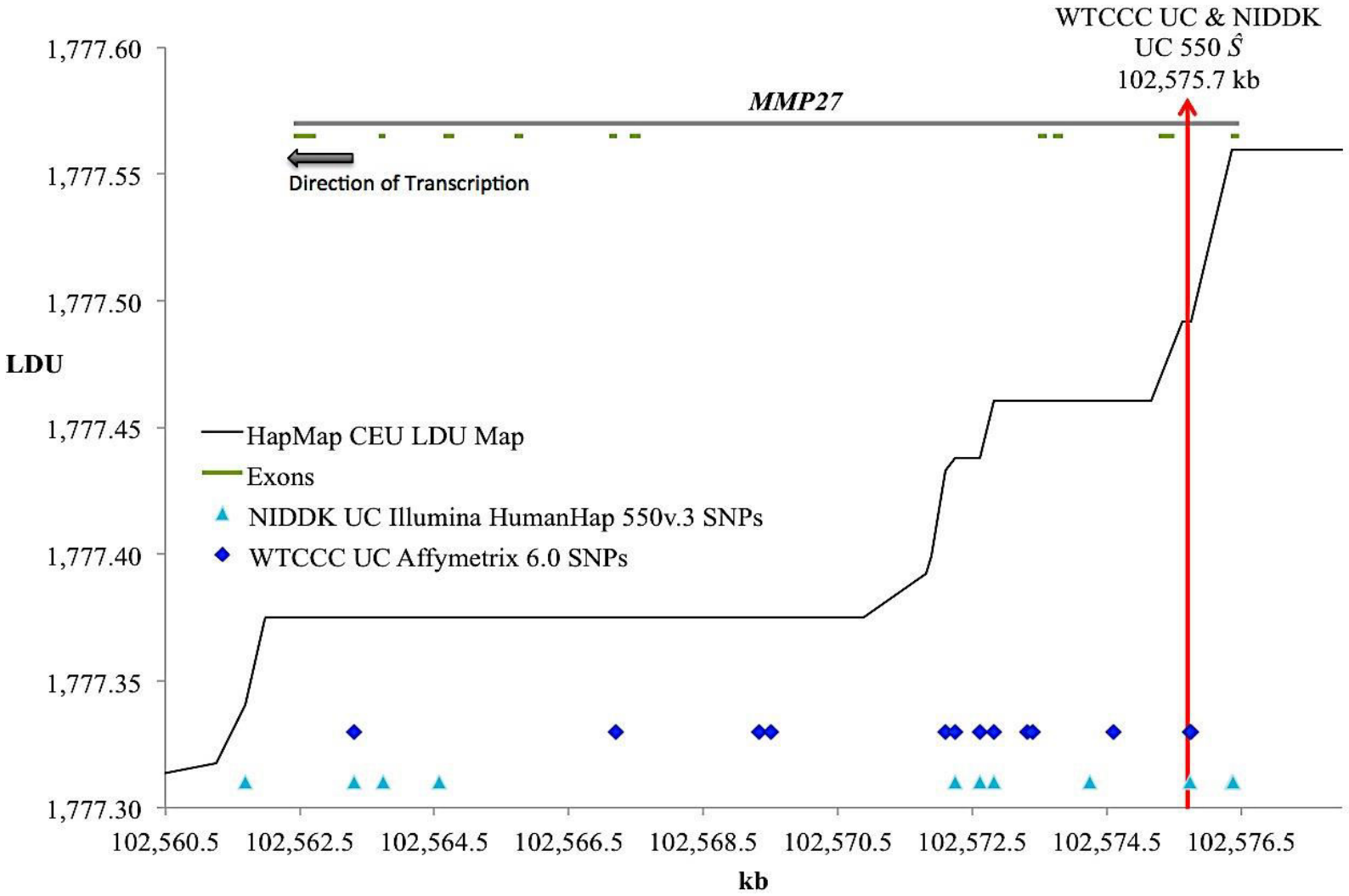

WTCCC UC & NIDDK
UC 550 Ŝ
102,575.7 kb
MMP27
Direction of Transcription
1,777.60
1,777.55
1,777.50
LDU
1,777.45
1,777.40
1,777.35
1,777.30
HapMap CEU LDU Map
Exons
NIDDK UC Illumina HumanHap 550v.3 SNPs
WTCCC UC Affymetrix 6.0 SNPs
102,560.5
102,562.5
102,564.5
102,566.5
102,568.5
102,570.5
102,572.5
102,574.5
102,576.5
kb

Figure 28. Localisation within the *MMP27* gene-region.
The Y-axis represents genomic LDU locations whereas the X-axis represents genomic physical locations in kilobases. The solid black line represents the LDU map of the region constructed from the HapMap CEU data. The red vertical arrow is the estimated location $\hat{S}$ of the putative causal agents. The dark green lines highlight the exons of the *MMP27* gene. The light blue triangles show the SNPs found in this region on the NIDDK UC Illumina HumanHap 550v.3 array and the dark blue diamonds are the SNPs on the WTCCC UC Affymetrix 6.0 array in this region.

IV. DISCUSSION

In this Chapter, the LDU multi-marker mapping approach successfully identified, with the additional provision of location estimates rather than LD intervals, 53 out of 59 previously reported UC loci. There could be several reasons for the failure in replicating the six remaining previously identified intervals. One reason could be the discrepancy in SNP resolution between the study carried out here and the meta-analysis that originally identified these intervals, since imputation was used in the meta-anaysis, which greatly increased the SNP resolution. Additionally, the original meta-analysis contained six GWAS datasets and therefore there is missing SNP data in the study carried out in this chapter, which solely made use of the WTCCC and NIDDK UC data. However, imputation was not carried out for the UC GWAS described in this chapter. Additionally, their efforts to combine data for the purpose of power may have increased heterogeneity within the dataset. Unlike in the case of the NIDDK CD dataset, there is lack of information on the UC phenotype. Based on the results form Chapters II and III, we can only predict that this increase in heterogeneity may translate to missing information.

In addition, 138 novel gene-regions were identified, with a closer replication of location estimates between the UC datasets. For the previously identified intervals, most of the signals were identified using the WTCCC UC data alone using this method and thus were not meta-analysed. The previously reported intervals together with the novel loci identified in this project increases the

number of UC loci found to date to 197, which is quite a substantial number considering that although polygenic, UC has been judged to have a lower genetic contribution than CD. It could be argued that even if some of these signals are false positives, many of the putative causal agents involved in UC are thus likely to be of low effect size. However, in the presence of disease heterogeneity, some of the signals could have a larger effect size in fewer people.

Previous genetic and animal studies have provided evidence for the involvement of a defective epithelial barrier in the gut (Consortium et al., 2009; Gitter et al., 2001; Schmitz et al., 1999). In the current study, several genes involved in cell motility, cell-cell junction and cell adhesion were identified. Examples of such genes are *CTNNA2*, *TLN2*, *FLRT2*, *FLRT3* and *PKP2*. Box 2 provides functional information on selected genes from Tables 13 and 14 a) and b). The molecular and functional Gene Ontology terms attributed to all the novel genes that were identified in this Chapter are given in Appendix V.

It is also interesting that some of the genes identified here are involved in peptidase activity, such as *ERAP1*, *USP3*, *USP35* and *TRHDE*. *ERAP1* (Endoplasmic Reticulum AminoPeptidase 1), illustrated in Figure 27, is a zinc metallopeptidase known to be implicated in the processing of antigenic peptides prior to presentation on the Major Histocompatibility Complex (MHC) Class I molecules (Agrawal and Brown, 2014). It is also involved in the regulation of the adaptive immune response and the inhibition of *ERAP1* has been found to increase the presentation of cell surface antigens as well as that of cytotoxic T-

cell responses (Zervoudi et al., 2013). *ERAP1* has also been previously identified as a general IBD susceptibility locus in the latest IBD meta-analysis, with the reported SNP located approximately 100 kb downstream of the estimated locations presented here (Jostins et al., 2012).

Some of the genes that encode proteins with peptidase activity also form part of the proteinaceous Extra-Cellular Matrix (ECM), for example *MMP27* and *NAV2*. For instance, *MMP27* is a very interesting candidate that has been found for the first time here, with identical estimated $\hat{S}$ locations between the WTCCC UC and NIDDK Illumina HumanHap 500v.3 data, located in intron 17 of the *MMP27* gene. *MMP27* is part of the Matrix MetalloProteinase (MMP) family (Figure 28).

MMPs are involved in the degradation of components of the ECM, such as collagen, which is an important process that contributes to processes such as cell survival, cell growth, as well as cell migration and cell invasion (Noel et al., 2012). MMPs are secreted as inactive proteins, which then become activated after cleavage by extracellular proteinases. Proteinases like MMPs disrupt the physiological barriers in order to facilitate the migration of cells and the release of chemotactic and growth factors from the ECM. Dyresgulation of MMPs contributes to cancer, as well as ulcer formation, which is central to the phenotype of UC (Nagase et al., 2006). Dysregulation of MMPs has also been identified in coronary pathology, such as Myocardial Infarction (MI) and Coronary Artery Disease (CAD), where both conditions exhibit inflammatory features. However, although the chicken MMP27 protein has been found to digest gelatin

and casein and causes autolysis of the MMP27 enzyme, little information is available on the mammalian MMP27 enzyme (Nagase et al., 2006).

Another interesting candidate identified in this part of the project is the Ubiquitin Specific Protease *USP3*, which is a hydrolase enzyme that deubiquitinates monoubiquitinated target proteins. Two of its target proteins are H2A and H2B, which are proteins involved in DNA-Damage Repair (DDR) check-point of the mitotic cycle. Thus, USP3 is essential for the proper progression through S phase and subsequent mitotic entry (Nicassio et al., 2007). As mentioned in the introduction of this chapter, studies have shown that improper cell-cycling is exhibited in UC patients. Indeed, it was found that the cell-cycle of colonic cells in UC patients progresses faster than that of colonic cells in healthy controls, which suggests a role for a dysregulated cell-cycle in UC and possibly why the progression to colorectal cancer is frequent among UC patients. *USP3* has also been found to be associated with both CD and UC in an association study that targeted genes encoding proteases that also identified *CYLD*, as discussed in Chapter II (Cleynen et al., 2014).

Other genes involved in the cell cycle have been identified in the UC meta-analysis described here, namely *MPHOSPH6*, which is a protein phosphorylated in the M-Phase of the cell cycle, and *PPP3CA,* a Serine/Threonine protein phosphatase, thus similar to cyclin-dependent kinases (cdks) and *RGS14*, which has also been found to be required for phagocytosis (Lim et al., 2013).

In summary, using the LDU mapping approach, many novel signals have been identified, within or close to genes that are functionally relevant to the UC phenotype, as well as that of colorectal cancer. However, it is important to note that the genes presented here are the genes which are closest to the location estimate obtained and therefore are not necessarily causal since long-acting cis elements may be involved. Also, promoters and enhancers located within one gene but regulate a more distant gene are also known to exist and this cannot be identified with the LDU mapping approach or other association testing methods.

Box 2: Functional overview of some genes of interest identified in Chapter IV [Tables 13 and 14 a) and b)].

***TRHDE*:** (Thyrotropin-Releasing Hormone-Degrading Enzyme). TRHDE is an extracellular peptidase, a member of the peptidase M1 family that specifically cleaves and inactivates the neuropeptide thyrotropin-releasing hormone. This gene was also identified in a GWA study of Rheumatoid Arthritis and in a study on Grave's Disease where both diseases display a huge involvement of a dysregulated immune system.

***CTNNA2*:** (Catenin, alpha-2). *CTNNA2* is also known as Cadherin-associated protein as a result of the physical interaction between *CTNNA2* and Cadherins. Catenins anchor Cadherins to the cell's cytoskeleton. Together with *CDH1*, which encodes E-Cadherin (See Chapter II), *CTNNA2* was found to be a mutator driver gene involved in gastric cancer (Wang et al., 2014).

***ERAP1*:** (Endoplasmic Reticulum Amino-Peptidase 1). Like TRHDE, ERAP1 is a member of the M1 family of zinc metalloproteinases. *ERAP1* is involved in the processing of immune-related receptor ligands, such as HLA-B27 ligands, via peptide trimming. Upon LPS stimulation, it has been demonstrated that the secretion of ERAP1 is mediated by Toll-Like Receptors (TLRs) through intermediate cytokines such as IFN-β and TNF-α that ultimately results in the activation of processes that mediate inflammation. *ERAP1* was also found to be associated with Ankylosing Spondylitis, Psoriatic Arthritis and Bechet's Disease, which are all conditions with an inflammatory aetiology.

***PKP2*:** (Plakophilin 2). PKP2 is a protein located in the desmosomes, which are cellular structures that link cells together, as well as in the cell nucleus. PKP2 is essential in bridging the cytoplasmic end of Cadherin proteins, such as E-Cadherin, with cytoskeletal intermediate filaments. PKP2 may be involved in the re-organisation of actin leading to the assembly of desmosomes.

CHAPTER V. GENERAL DISCUSSION AND FUTURE WORK.

i. The story so far.

Disease-gene mapping using linkage analyses has a long history, but association mapping was still emerging when techniques that allow high-throughput SNP genotyping were being developed. The success of GWAS as a method of gene-mapping has been surrounded by skepticism from much of the scientific community since although many loci have been identified, a substantial number of cases and controls (75,000 cases and controls combined in the latest IBD meta-analysis) were needed to achieve these results and the genetic contribution that these loci explain remains low. Another factor that contributed to this skepticism is the fact that the genes that have been identified in GWAS have been done so by *in silico* analyses of LD intervals. Indeed, in the case of large LD intervals or LD-intervals which are gene-rich, many candidate genes have been reported and thus it is not known which are the actual putative causal agents within these intervals.

a. Genetic and phenotypic heterogenetiy.

In this project it was firstly shown that with detailed phenotypic information, one can detect a greater proportion of the genes involved, as well as phenotypic and genetic heterogeneity, even with the use of smaller data sets, and thus sample sizes. By scanning the genome using a method based on composite likelihood that takes directly into account the underlying LD structure in the genome, more information than previously reported was extracted from these publically available GWAS datasets without recourse to enormous meta-analyses. In addition, the estimated locations of the putative causal agents that this LDU mapping method provides is essential in narrowing down the locations of the potential putative causal agents. The location estimates provide a starting point for targeted functional analyses and sequencing. Another major advantage of location estimates of the putative causal agents is the ability to more easily identify the possible presence of genetic heterogeneity, which is known to be involved in the aetiology of many complex diseases (Silverman and Palmer, 2000). Previous studies by our group and others have already reported the increased power of using composite likelihood, rather than single-SNP testing and modelling on LDU distances as opposed to physical locations in kb (Andrew et al., 2008; Maniatis et al., 2007; Politopoulos et al., 2011). The work presented in this thesis, part of which is published, illustrates this further by extending to genome-wide data (Elding et al., 2011, 2013).

In Chapters III and IV a large number of signals associated with CD and UC were identified, which mainly identify single genes with confirmation in independent cohorts, making it likely that most are real. In some cases they may have a large effect and in some cases only a small effect on the risk of CD or UC. Since the location of the causal variant is based on an estimation within a 95% confidence interval, the actual causal variant or variants that are in direct LD are still to be found and it is thus not possible to measure the effect size or indeed the direction of effect of the causal variant. Critically, as also demonstrated in the second chapter, it is likely that different risk genes are involved to a greater or lesser extent in different patients. The genomic locations that are provided in this project are estimates of the positions of the variants of functional significance assuming a mono-phylogenetic origin and thus no substantial allelic heterogeneity. Previous studies demonstrated the greater power of using LDU as compared to physical maps for detecting association and refining location using the same multi-marker methodology (Andrew et al., 2008; Gibson et al., 2008; Politopoulos et al., 2011). The close distance between the location estimates for many of the signals in the two different cohorts analysed in Chapter III and IV suggests that these allelic variants will often be the same, but in some cases there are clear indications of independent allelic variants, perhaps controlling the expression of the same gene (in *cis*). Genetic heterogeneity can also be attributed to unlinked loci.

It has been previously found that the reported *NOD2* mutations determine ileal disease only and that disease heterogeneity may be genetically determined, coupled with the fact that ileal disease is generally associated with an early age

of diagnosis (Ahmad et al., 2002). This therefore demonstrates the importance of considering phenotypic subgroups of CD and hence collecting more accurate phenotypic description of patient data for GWAS, especially since it is well established that response to treatment in CD patients is somewhat dependent on the disease location rather than disease behaviour (Weiss et al., 2010). The 95% CIs of the estimated locations $\hat{S}$ can help in pin-pointing such cases where independent allelic variants are likely to play a role. An example is the case of *IRF8* described in Chapter II (Figure 9) where although the SNP coverage in the region is quite similar between the two datasets, the estimated locations have non-overlapping 95% CIs. Although this could be attributable to the only partially-overlapping genotyping arrays used in both studies, it could also be an indication of the involvement of independent allelic variants at that locus. In either case this method offers greater precision for future in-depth analysis using bioinformatics and re-sequencing.

One major problem is that phenotypic heterogeneity is currently under-studied, especially in large-scale genetic studies. For instance, the prevalence of CD among Ashkenazi Jews is higher than that for non Jewish Europeans, but surprisingly, in 1996 it was shown that no linkage is observed on chromosome 16q in Ashkenazi Jewish families (Ohmen et al., 1996). Also, as discussed in Chapter I, no linkage was ever observed on chromosome 16q in UC families (Ohmen et al., 1996), which was one of the first lines of evidence for the presence of genetic heterogeneity between the two sub-types of IBD. Additionally, although there is a substantial overlap in susceptibility loci between early-onset and adult-onset, there is also a degree of heterogeneity. For instance,

the signal on Chromsome 16p (Table 8 and 15) was found to be associated in early-onset CD patients (Imielinski et al., 2009). Indeed, in this thesis work, the signal on Chromosome 16p was identified using the NIDDK CD dataset only (Table 8 and 15), which contains patients with early-onset CD, but not in WTCCC. Despite this knowledge, disease phenotype is not well recorded in the publically available data-bases and heterogeneity for neither CD or UC has been studied in detail to date. The focus has been toward large meta-analyses combining different sub-types and age of onset.

The main reason for the limited study of heterogeneity is because the main concern has been to increase the sample size and hence power to detect association. The extent to which the meta-analysis approach results in a gain in power is variable since combination of groups of patients who have been ascertained by different criteria and from different geographic locations will mean that variants that only apply to a specific sub-group will go unseen. Here I show that stratifying the data based on subtypes of CD proves to be more beneficial in dissecting the underlying genetic heterogeneity involved in CD, in contrast to increasing the sample size, at least with the method reported here. It will be important in the future to reconsider the collection of clinical data and attempt to record more complete information.

The previous linkage peaks significant for IBD (Cavanaugh, 2001; Cavanaugh et al., 1998; Hampe et al., 1999; Hampe et al., 1999; Hugot et al., 1996; Ohmen et al., 1996; Parkes et al., 2000; Satsangi et al., 1996) have all been identified in

GWAS, with the exception of *IBD4* in relation to CD and *IBD6* on Chromosome 19p13 in relation to UC, as shown in Table 15. In the linkage study that identified *IBD6*, it was reported that there was substantial heterogeneity in relation to CD with respect to significance of linkage and the carrier status of the *NOD2* mutations in CD patients. The same study did not find significant linkage at the *IBD4* locus (van Heel et al., 2003). Indeed, the *IBD6* locus was only identified after the patients with the *NOD2* mutations were removed from the data, suggesting the presence of genetic heterogeneity. Since the same study detected evidence for the presence of epistatic interactions between the *IBD5* and the *IBD6* loci, the authors hypothesized that the Canadian study that detected linkage at the *IBD6 l*ocus could be explained by the fact that their study showed a stronger *IBD5* linkage signal than the one reported by van Heel *et al*. and could thus facilitate the detection of linkage at the *IBD6* locus in the Canadian families. The *IBD8* UC linkage peak is identified in a GWAS for the first time in the work described in Chapter IV of this thesis.

Table 15. Follow up of significant IBD linkage peaks in GWAS.

			CD		UC	
Locus	**Chr**	**Linkage Significance**	**Confirmed from previous Meta-Analyses**	**Newly reported in this thesis**	**Confirmed from previous Meta-Analyses**	**Newly reported in this thesis**
IBD1	16q12	CD	*NOD2*	*CYLD*	-	-
IBD2	12p13.2-q24.1	CD and UC	*LRRK2/SLC2A13*	*ANO4* *C12orf74* *TMEM117* *CCDC91* *CAND1* *KITLG*	*IFNG*	*CCDC91* *PKP2* *TRHDE* *WSCD2/CMKLR1*
IBD3	6p21.3	CD and UC	*HLA-DQB2*	-	*TAP2/HLA-DQB-2*	*HLA-DPB1/HLA-DOA*
IBD4	14q11-q12	CD	***Not identified in GWAS***	-	*Not identified in GWAS*	*NOVA1/MIR4307*
IBD5	5q31	CD	*C5orf56/SLC22A4*	-	*C5orf66*	*HSPA4/FSTL4* *NRG2* *YIPF5*
IBD6	19p13	CD and UC	*SBNO2* *PDE4A*	-	***Not identified in GWAS***	-
IBD7	1p36	CD and UC	*CAMTA1*	-	*OTUD3/RNF186* *ZBTB40*	*CATSPER4*
IBD8	16p	CD and UC	*CD19*	*RBFOX1*	***Not identified in GWAS***	*GPR139*

Genes separated by "/" represent one signal from the same window but estimated locations $\hat{S}$ of the causal variant for the discovery and replication dataset(s) point to different genes. Red font implies that although linkage has been observed, no signal has been detected in previous GWAS or in this thesis work. Blue font represents significant linkage peaks that have not been detected by previous GWAS but association in the linkage region has been identified in this thesis work. The significance and estimated locations $\hat{S}$ for these signals are illustrated in Tables 8, 9 and 10 for CD and in Tables 12, 13 and 14 for UC.

It is also of interest to consider our GWAS results described in Chapters III and IV in relation to previous candidate gene studies. Like single-SNP analysis, the LDU mapping approach still failed to replicate some genes that had been reported to be associated from single gene studies. For instance *IL16*, *TLR4* and *MUC3A* (Franchimont et al., 2004; Glas et al., 2003; Kyo et al., 2001; Ouburg et al., 2005) are among the examples of genes that were good functional candidates in early candidate gene studies but were never identified in GWA studies. In the case of *IL16*, functional analysis revealed the involvement of a polymorphism located in promoter region of *IL16* in CD patients only (Glas et al., 2003) and an amino acid substitution (Asp299Gly) polymorphism in the *TLR4* gene was reported to be involved in both CD and UC (Franchimont et al., 2004). Apart from possible ethnic or disease sub-group differences, the failure to replicate these loci could be simply due to the fact that these loci might have been false-positives in the previous studies. Another reason for failing to replicate true susceptibility loci in functionally implicated genes could be the low SNP coverage or problems with these regions affecting the current genotyping platforms. In the case of *MUC3*, the Genome Browser sequence is still incorrectly annotated and fails to detect that it is a duplicated gene (Gum et al., 2003; Kyo et al., 2001; Pratt et al., 2000), suggesting that the LDU map is also inaccurate, since an error in the physical map will be reflected in the genetic map.

b. Interactions and Epistatic effects.

Aside from the impact of accurate phenotypic description and that of heterogeneity, another complication in identifying susceptibility loci for IBD, or complex diseases in general, arises from the presence of interactions or epistasis between loci. Indeed, this could be another reason behind the failure in replicating findings from functional candidate gene studies in association studies and meta-analyses. *DMBT1,* a scavenger and antibacterial pattern recognition receptor, was never identified in GWASes and is an example of such interaction. A candidate gene study had identified an association between a deletion variant of the *DMBT1* allele and and increased risk for CD (Renner et al., 2007). More recently functional and epistatic analyses have shown evidence of interaction. *DMBT1* is a target gene of the *IL-22* cytokine, which is in turn modulated by CD-risk variants in *IL23R*, an established IBD susceptibility locus (Zervoudi et al., 2013). The risk variants in *IL23R* have also been shown to modulate the expression of *DMBT1* in the intestines of CD patients (Zervoudi et al., 2013).

Another example comes from *NOD2*. As mentioned in the introduction of Chapter IV, it had already been established that *NOD2* is a CD-specific risk factor but although linkage and association analyses failed to identify *NOD2* in relation to UC, it was still shown that there is a possible epistatic interaction between the cytokine cluster-containing *IBD5* locus and the *NOD2* mutations and that these two loci could interact in a synergistic way to contribute towards the onset of UC (Giallourakis et al., 2003). In fact, the same study also reported that

much of the significance of the *IBD5* association signal arose from UC patients that carried the *NOD2* mutations (Giallourakis et al., 2003).

In addition to epistatic effects, gene-environment interactions are of great importance in complex conditions and can have a considerable impact on the development of the phenotype. In fact, an individual can inherit the predisposing factors to a particular disease but never develop the disease itself unless exposed to an environmental factor. Therefore, the analogy that "genetics loads the gun but the environment pulls the trigger" perfectly captures the complex interplay between host genetics, including gene-gene-interactions, and disease onset (Olden, 2006; Ramos and Olden, 2008). After all, autoimmune or immune deficiency diseases like IBD involve the host's response to an environmental stimulus. Thus, in a genetically susceptible host, this gene-environment interaction can lead to the onset of disease. Indeed, the term gene-environment interaction can be defined as the environment's impact on the direction and degree of a genetic variant's effect on phenotype (Ramos and Olden, 2008). In the case of IBD, it is well established that diet and the gut microbiota composition of an individual are major environmental factors that play a role in the onset of IBD. In addition, smoking is an environmental factor which shows different directions of effect with respect to CD and UC, with smoking apparently being 'protective' for UC. Gene-environment interaction studies have so far been hampered by the fact that studies in experimental models they have to be done in model animals or cell lines and are usually limited to one environmental component, which may not be sufficient to show what occurs *in vivo.*

Further epidemiological and genetic association studies that contain data on environmental factors are needed. These will provide information that will eventually help in elucidating the mechanisms and pathways involved in disease onset as well as in providing the opportunity for possible intervention prior to disease onset, where predisposing genetic factors are known to be involved.

ii. Considerations for future work involving the LDU mapping approach

Although the LDU mapping method used in this project shows several advantages over other types of association methods, there are limitations that need to be considered. As for the case with other association testing methods, this method is not immune to false-positive results. A Bonferroni threshold was used throughout this project in an attempt to account for false-positives arising purely by chance due to multiple testing. Bonferroni correction is known to be a very conservative method to account for multiple testing, which makes it a difficult case to find a fine balance between eliminating false-positives and reducing false-negatives as a result of multiple testing corrections in association studies. However, false-positives could also arise due to the presence of population sub-structure in the datasets. For all the datasets that were analysed in this project, the data were projected on the first two Principal Components of HapMap3 data in order to detect any genome-wide population sub-structure within the GWAS

data. This method for detecting structure is routinely done in association studies and is thought to work well on a genome-wide level. Indeed, in the quality control analyses of previously published GWAS, as well as in the quality control that was carried out in this thesis work population outliers that were known to exist within the data were successfully identified. However, what if the data exhibits hidden sub-structure, such as locus-specific effects which reflect very subtle differences in recent ancestry?

In recently admixed populations, such as in African Americans, two or more populations would have been mixing for a short amount of time thus leading to a new population (Glas et al., 2003). The ancestry of this new population can be accounted for by the different proportions from the original populations. Unlike in non-admixed populations, where the whole genome is representative of the population under study, in these recently admixed populations the genome is subdivided, due to recombination, into short sections of differing ancestry, which makes it difficult to identify the ancestral genetic information of such individuals (Glas et al., 2003). Although the scenario of such admixture is highly unlikely to apply to the datatsets used in this thesis work, independent work carried out on UK samples as part of the People of the British Isles (PoBI) project identified the presence of population structure within the UK when surnames were taken into account and the data were modeled using a simple point admixture approach based on maximum likelihood (Winney et al., 2012). This fine population structure within the UK was never detected when traditional methods, such as PCA or Fst, were carried out. However subtle stratification due to difference in the representation of geographic regions in the UK is masked in the PCA since there

are only short segments of the genome that differ, in other words there is the presence of locus specific differences in ancestry. One of the factors that contributes to these regions of local ancestry is natural selection. Indeed, IBD is an immune-mediated disease and natural selection has played a significant role in shaping host immune defence by acting on immune-related genes (Barreiro and Quintana-Murci, 2010). Natural selection on immune related genes, whether negative, positive or balancing selection, reflects past infectious epidemics and thus past population demographics. The effect of this selection on host immunity could hence partly contribute to the differences in prevalence of certain diseases in different populations (McEvoy et al., 2009). Locus specific ancestry differences could thus be one of the causes of spurious association with any association testing methods used in complex disease genetics.

Another issue to be considered when using the LDU mapping method is the selection of the difference in distance between the estimated locations $\hat{S}$ between datasets in order for a location to be considered a replication. In fact, for the main CD meta-analysis carried out in Chapter III the difference in distance between the estimated locations $\hat{S}$, for the locations where the meta P-value survived the Bonferroni threshold, was set to lie within 150 kb between the WTCCC CD and NIDDK CD datasets, but the majority fell within 80kb. Since the UC SNP data had a much higher SNP resolution (1 million SNPs) I used the more stringent estimated location $\hat{S}$ difference threshold of 80 kb. However, making the decision about this difference threshold is rather arbitrary. Another approach would be to use a LDU cut-off for distance difference thresholds or to take into account the 95% Confidence Intervals of the estimated locations $\hat{S}$. Clearly, the difference in

SNP content and resolution in the different datasets that are meta-analysed may affect the estimations of $\hat{S}$. A computational approach to resolving this dilemma would be complicated.

This highlights the problem of the heterogeneity introduced into the datasets by the difference in SNP coverage on the platforms used at different times by the various investigators. A solution that investigators have used to overcome this problem has been to impute genotypes in order to homogenise the set of SNPs that are being tested in the different datasets. Imputation could also potentially be used in conjunction with the LDU mapping approach, in order to meta-analyse the same set of SNPs across datasets. However, the advantage of the LDU mapping approach is that comparisons can be made across data sets even with very different SNP coverage.

A high-resolution, homogeneous subset of SNPs across datasets, together with better refined phenotypes should reduce the complications that heterogeneity brings to disease-gene mapping.

iii. Future work- Inflammatory Bowel Disease meta-analysis and beyond.

The work that was carried out during the course of this thesis focused on CD and UC specific analyses. The next step would be to conduct an analysis specifically for IBD. There are two possible approaches that could be taken in

order to carry out this work. The first method would involve the meta-analysis of the location estimates that are close to each other from the nominally significant windows in the CD and UC datasets independently and subsequently meta-analysing the P-values and applying a Bonferroni correction for multiple testing. Taking this approach will involve the careful consideration about what difference in distance between the estimated locations obtained by the CD and UC datasets should be used, especially in the light of the considerations that have been discussed above, since genetic heterogeneity involving the same genetic region will become more likely when considering CD and UC together.

The second approach would be to combine the genotype data for both diseases, for example combining the WTCCC CD with WTCCC UC as one data set and combining the NIDDK CD together with the NIDDK UC datasets as a replication data set. The major problem with this approach is that all the CD and UC datasets are genotyped on different genotyping platforms. Therefore the direct concatenation of datasets will not be possible, and SNPs that are not common to the genotyping platforms of the data being concatenated will be lost, thus leading to a loss in SNP resolution. However, these SNPs could perhaps be imputed, without having a huge impact on the LDU mapping approach since much of the power of this approach comes from the LDU maps, which are obtained using the HapMap data. This approach would also overcome the issue that the NIDDK CD Controls would need to be used twice if the UC and CD datasets were to be analysed separately, since the NIDDK CD Controls were used in Chapters II, III and IV. Indeed, both approaches could be carried out and the results can be compared, although in my opinion, the second approach would

be best as it would require less decision-making about distance threshold for estimated locations that could affect the final results and it also provides the opportunity to increase the sample size for each CD-UC dataset, which will lead to an increase in power to detect association pertaining to IBD, given that in this case only loci that are in common between the disease categories are of interest. Thus, the second approach seems to provide a better balance between Type I (False Positives) and Type II errors (False Negatives).

In the meantime, it is worthwhile to look at the genes that have been identified in both disease sub-types independently, in the work described in Chapters III and IV. Table 16 illustrates the genes that are common to both CD and UC and were found independently in Chapter III (Tables 8 and 9) and Chapter IV (Table 14).

Table 16. Genes common to both CD and UC identified in this thesis work.

Chr	LD Interval (Mb)[§]	WTCCC CD $\hat{S}$	NIDDK CD $\hat{S}$	WTCCC UC $\hat{S}$	NIDDK UC 550 $\hat{S}$	NIDDK UC 300 $\hat{S}$	Overlap in 95% CI of significant estimated locations $\hat{S}$	Gene
1p	67.53-67.77	67,864.8	67,707.3	67,769.1	-	67,741.5	WTCCC CD/NIDDK CD/NIDDK UC 300	*IL23R**
2q	-	170,038.8	170,037.2	169,969.0	170,001.5	-	None	*LRP2**
2q	-	226,258.2	226,258.4	226,257.7	226,231.6	-	WTCCC CD/NIDDK CD/WTCCC UC	*NYAP2*
5p	40.28-40.81	40,447.1	40,288.0	-	40,617.6	-	WTCCC CD/NIDDK UC 550	*PTGER4*
5q	158.53-158.93	158,825.7	158,826.0	158,825.8	-	-	All	*IL12B*
7p	-	41,898.3	41,939.6	41,923.9	-	41,941.0	All	*GLI3*
9q	117.44-117.70	117,555.6	-	117,559.8	-	-	All	*TNFSF15**
10p	35.18-35.90	35,554.8	-	35,554.4	-	-	All	*CCNY**
10q	-	63,604.2	63.553.3	63,552.9	63,553.5	-	NIDDK CD/WTCCC UC/ NIDDK UC 550	*C10orf107*
10q	-	90,787.7	90,787.7	90,739.8	90,787.4	-	WTCCC CD/NIDDK CD/NIDDK UC 550	*FAS*
10q	101.26-101.34	101,324.7	101,274.6	101,754.7	101,342.5	-	None	*NKX2-3*
11q	76.04-76.34	76,304.4	-	76,303.3	-	-	All	*C11orf30*
12p	-	28,744.32	28,741.7	28,6769.5	28,595.9	28,309.1	NIDDK CD/WTCCC UC/NIDDK UC 300	*CCDC91**
16q	-	85,982.3	85,958.0	85,996.3	-	-	WTCCC CD/WTCCC UC	*IRF8**
20p	-	15,802.7	15,876.4	14,314.2	14,352.4	-	None	*MACROD2**
20q	62.19-62.51	62,327.9	-	62,326.9	-	-	All	*RTEL1**
21q	16.70-16.85	16,841.4	-	16,818.4	16,802.9	-	All	*USP25*
21q	45.59-45.70	45,609.1	-	45,608.8	-	-	All	*ICOSLG*

* denotes that at least one of the estimated locations $\hat{S}$ is located within the identified gene. [§] LD interval if the signal was reported in the published meta-analysis previous to this thesis work. – means no significant signal detected within the datasets. Signals highlighted in green represent the signals that have been identified in this thesis work.

In the CD analysis, approximately half of the signals that were identified in this thesis work, highlighted in dark green in Table 16, were replicated using the smaller NIDDK CD dataset containing Ileal patients that exhibited extra-ileal manifestations and half using the NIDDK Pooled dataset (Table 9 and Table 10). In 5 of the 7 shared genes shown in Table 16, the difference in location estimates between both the CD datasets and the UC datasets is less than 70kb, but not for the remaining two signals. In the case of *CCDC91,* shown in Figure 29, the difference in location estimates was larger than 400 kb but genetically, the location estimates lie close to one another, being within approximately one LDU of each other. *CCDC91* is an interesting finding since although not much in known about this gene, Ph.D. thesis work carried out by Tao Jiong at the National University of Singapore that focused on looking at genomic copy number alterations identified a deletion of the *CCDC91* gene region in a gastric cancer cell line (Jiong, 2011). In addition, independent studies have identified this region as a CNV region, where both amplification and deletion of the *CCDC91* gene and the surrounding region have been detected in cancer cell lines (Bignell et al., 2010; Johnson et al., 2010; Shih Ie et al., 2011). In the case of *MACROD2*, the location estimates between CD and UC were approximately 1,500 kb apart, but the location estimates for all the CD and UC datasets lie within the *MACROD2* gene. *MACROD2* is also located at a site of genomic instability, with recurrent microdeletions being reported within the *MACROD2* gene locus, both in the general population and in patients of colorectal carcinomas (Linnebacher et al., 2013). The genomic instability of these two regions, and possibly that of other regions along the genome, could explain the

difficulty in obtaining close location estimates between different datasets, since this instability is also very likely to be reflected in the LDU map of the region.

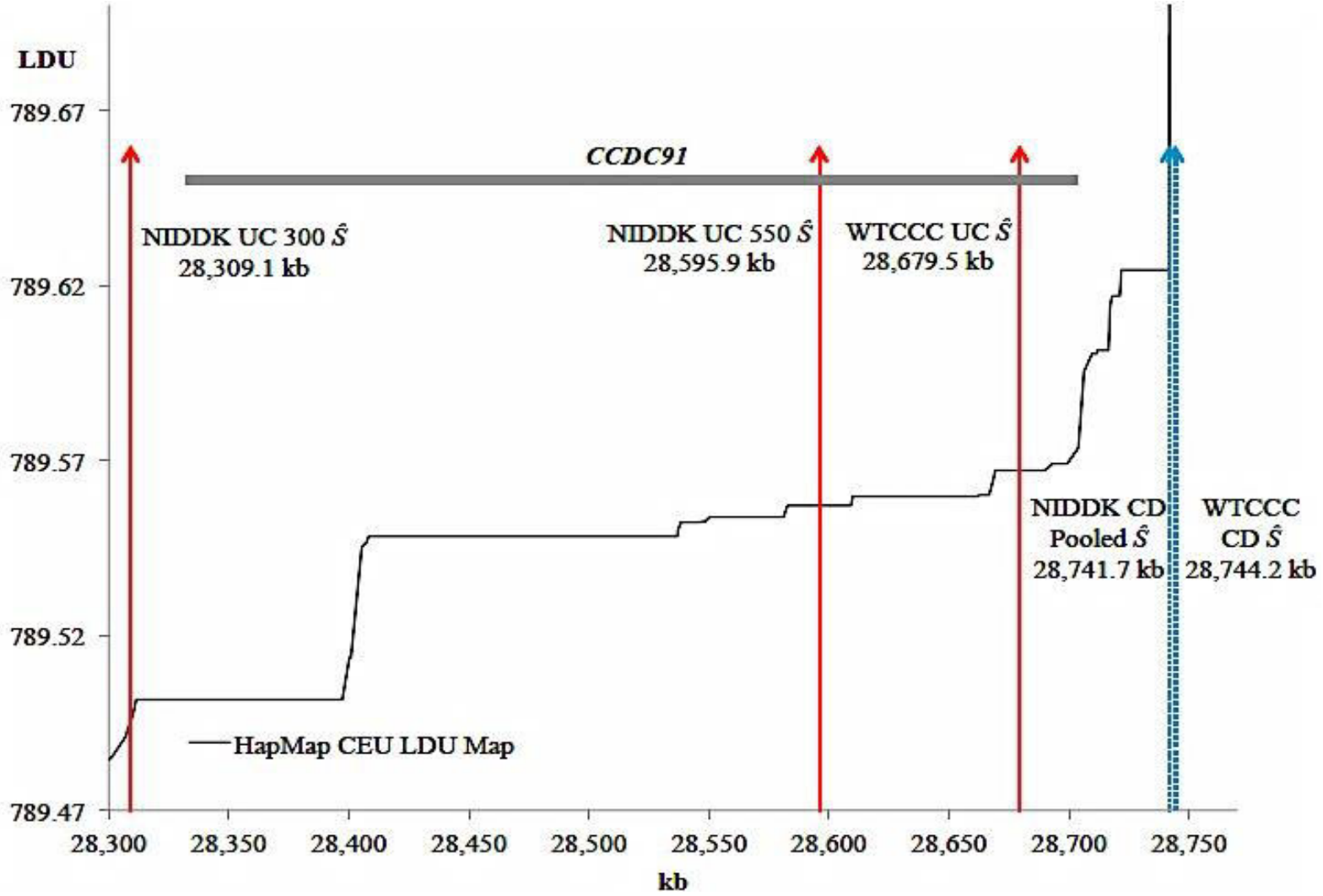

LDU
789.67
789.62
789.57
789.52
789.47
CCDC91
NIDDK UC 300 $\hat{S}$
28,309.1 kb
NIDDK UC 550 $\hat{S}$
28,595.9 kb
WTCCC UC $\hat{S}$
28,679.5 kb
NIDDK CD
Pooled $\hat{S}$
28,741.7 kb
WTCCC
CD $\hat{S}$
28,744.2 kb
HapMap CEU LDU Map
28,300
28,350
28,400
28,450
28,500
28,550
28,600
28,650
28,700
28,750
kb

Figure 29. Localisation within the *CCDC91* gene-region.
The Y-axis represents genomic LDU locations whereas the X-axis represents genomic physical locations in kilobases. The solid black line represents the LDU map of the region constructed from the HapMap CEU data. The red vertical arrows are the estimated locations $\hat{S}$ of the putative causal agents for the UC data. The blue vertical dashed arrows are the estimated locations $\hat{S}$ of the putative causal agents for the CD data.

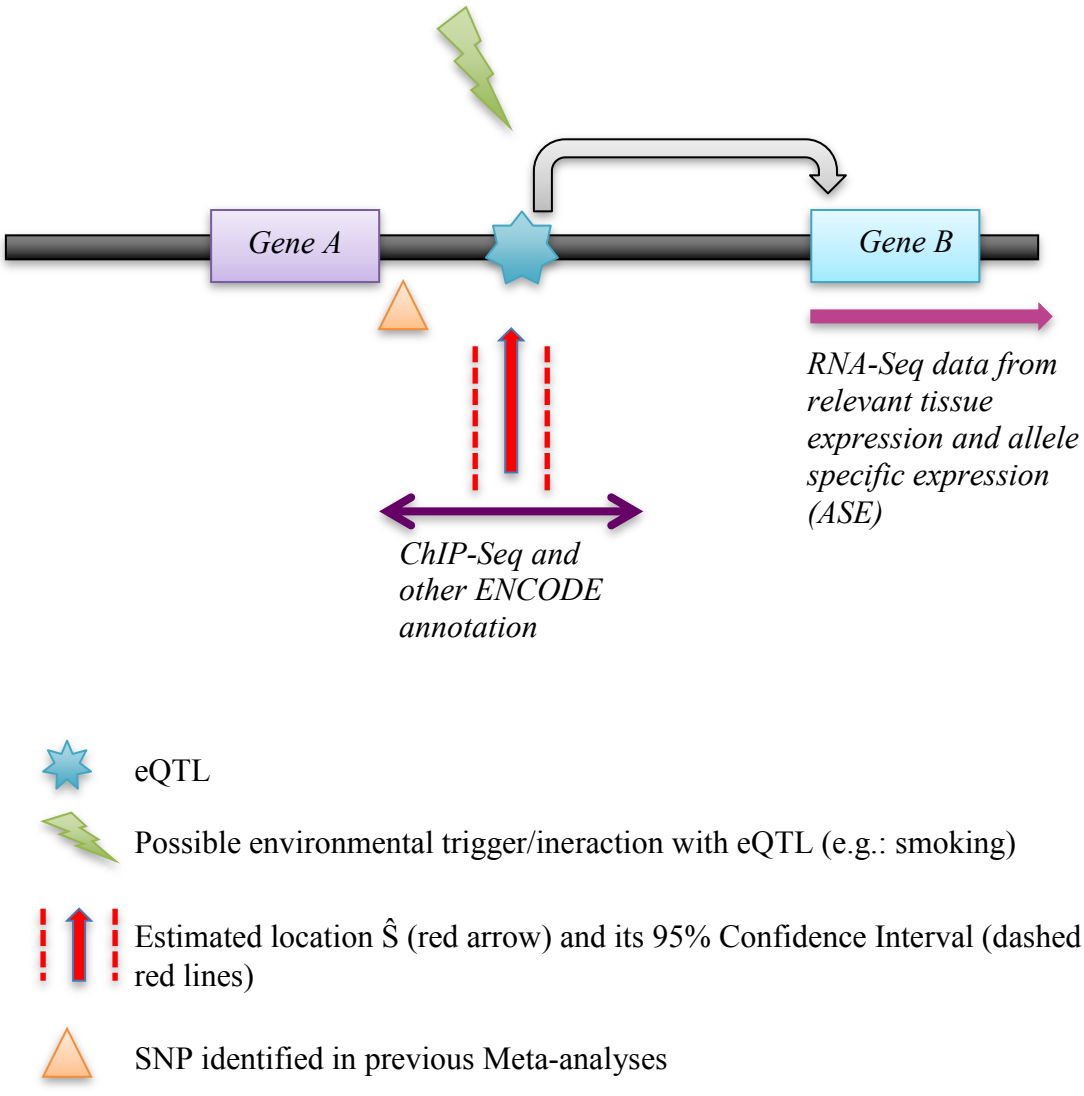
Other future work would also be to combine the results from GWAS with intestinal expression data.

The co-localisation of eQTL loci with some of the estimated locations outside exonic regions obtained in this project would help to narrow the region housing the likely causal variants involved in regulation. This co-localisation could also be done in conjunction with ChIP-Seq (ChIP-Sequencing) analysis, and other functional annotations from the ENCODE project, which can be used to point to DNA-protein binding sites, especially for the identified regions that do not harbour any genes and could thus represent non-coding regulatory SNPs. This co-localisation approach is illustrated in Figure 30. Technological advances in Next-Generation Sequencing paved the way to RNA-Seq (RNA-Sequencing), or "Whole-Transcriptome Shotgun Sequencing", which is one way of carrying out expression analysis. One application of RNA-Seq is to explore the RNA levels from a genome at a given time. This tissue-specific expression information, together with the results from previous meta-analyses and the estimated locations identified in this thesis work could be consolidated in order to pinpoint the genes involved in CD and UC.

Considering that complex diseases, such as IBD, result from the contribution of both environmental and genetic factors, investigating the impact of the interactions between environmental factors and regulatory SNPs will also be a crucial element in dissecting the environmental triggers that contribute to the onset of IBD as well as in identifying the mechansims and pathways that lead to

the disease phenotype. In fact, interactions between the environment and regulatory SNPs affecting Allele-Specific Expression (ASE) has already been demonstrated in a study that looked at monozygotic twins (Buil et al., 2014).

Figure 30. A co-localisation approach for future studies of possible regulatory regions.



Genetic variation in regulation does not necessarily involve the closest gene. ASE can be conducted on biopsy material from patients.

Another approach will be the investigation of gene-gene interactions, which as discussed in this Chapter, are known to contribute to the onset of IBD. For example the approach of stratifying the data by *NOD2* could be used to explore the effect on unlinked associations, perhaps by examining loci flagged in the literature or perhaps genome-wide.

From the studies carried out on the gut microbiota in relation to IBD pathogenesis no causal bacterial agent has been identified so far and it is still unclear whether the dysbiosis observed is the result of, or indeed precedes, the onset of IBD, leading to the chicken and the egg dilemma- which of the two comes first? The improvements in techniques for microbial community profiling, such as the advent of Next-Generation Sequencing, opened the doors to high-throughput analyses of genomes of entire communities, in other words Metagenomic analyses, including previously unculturable organisms. The aim of metagenomics analyses is to determine the role that the microbes play in human health as well as in the onset of diseases. Indeed, the Human Microbiome Project (HMP) was conceived, which started in 2007 with the aim of developing the necessary tools and datasets in order to study the role that microbes play in relation to human health and disease. The project is currently in its second and final phase generating datasets that integrate the biological properties from both microbiome and host from studies related to micobiome-associated diseases and is expected to end in 2015. Such datasets could be exploited in order to unravel the host-microbiota interactions that take place in the pathogenesis of IBD and elucidate whether microbiota dysregulation takes place prior to or after the onset of IBD. The data generated from such studies could be used as phenotypes in

GWA studies in order to identify possible association between a particular microbiota species and disease outcome. Another approach would be to assess the clustering of the microbiota composition of CD or UC patients according the specific sub-phenotypes (e.g.: ileal only disease for CD and proximal versus distal colonic inflammation in UC). Depending on the findings, this would also provide the opportunity for therapeutic development and possibly prevention of dysbiosis in genetically susceptible hosts.

From previous association studies, as well as from the work presented in this thesis, it can be concluded that with increasing bioinformatics and statistical advances, association studies are identifying many loci at a much faster pace than can be functionally tested, not only for IBD, but for complex diseases globally. It is only very recently that functional techniques such as RNA-Seq are scaling up to genome-wide level. Combining the ever increasing statistical tools together with the emerging functional techniques, including the human microbiome data, could thus open the door to unraveling the mechanisms behind IBD and ultimately target the relevant pathways in the specific CD and UC subgroups for sub-phenotype stratified treatment.

Bibliography

Abbott, D.W., Wilkins, A., Asara, J.M., and Cantley, L.C. (2004). The Crohn's disease protein, NOD2, requires RIP2 in order to induce ubiquitinylation of a novel site on NEMO. Curr Biol *14*, 2217-2227.

Agrawal, N., and Brown, M.A. (2014). Genetic associations and functional characterization of M1 aminopeptidases and immune-mediated diseases. Genes and immunity.

Ahmad, T., Armuzzi, A., Bunce, M., Mulcahy-Hawes, K., Marshall, S.E., Orchard, T.R., Crawshaw, J., Large, O., de Silva, A., Cook, J.T., et al. (2002). The molecular classification of the clinical manifestations of Crohn's disease. Gastroenterology *122*, 854-866.

Akhabir, L., and Sandford, A. (2010). Genetics of interleukin 1 receptor-like 1 in immune and inflammatory diseases. Current genomics *11*, 591-606.

Amundsen, S.S., Rundberg, J., Adamovic, S., Gudjonsdottir, A.H., Ascher, H., Ek, J., Nilsson, S., Lie, B.A., Naluai, A.T., and Sollid, L.M. (2010). Four novel coeliac disease regions replicated in an association study of a Swedish-Norwegian family cohort. Genes and immunity *11*, 79-86.

Anderson, C.A., Boucher, G., Lees, C.W., Franke, A., D'Amato, M., Taylor, K.D., Lee, J.C., Goyette, P., Imielinski, M., Latiano, A., 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.

Andrew, T., Maniatis, N., Carbonaro, F., Liew, S.H.M., Lau, W., Spector, T.D., and Hammond, C.J. (2008). Identification and Replication of Three Novel Myopia Common Susceptibility Gene Loci on Chromosome 3q26 using Linkage and Linkage Disequilibrium Mapping. PLoS Genet *4*, e1000220.

Asano, K., Matsushita, T., Umeno, J., Hosono, N., Takahashi, A., Kawaguchi, T., Matsumoto, T., Matsui, T., Kakuta, Y., Kinouchi, Y., et al. (2009). A genome-wide association study identifies three new susceptibility loci for ulcerative colitis in the Japanese population. Nat Genet *41*, 1325-1329.

Barreiro, L.B., and Quintana-Murci, L. (2010). From evolutionary genetics to human immunology: how selection shapes host defence genes. Nat Rev Genet *11*, 17-30.

Barrett, J.C., Hansoul, S., Nicolae, D.L., Cho, J.H., Duerr, R.H., Rioux, J.D., Brant, S.R., Silverberg, M.S., Taylor, K.D., Barmada, M.M., 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., Lee, J.C., Lees, C.W., Prescott, N.J., Anderson, C.A., Phillips, A., Wesley, E., Parnell, K., Zhang, H., Drummond, H., et al. (2009). Genome-wide

association study of ulcerative colitis identifies three new susceptibility loci, including the HNF4A region. Nat Genet *41*, 1330-1334.

Beaud, D., Tailliez, P., and Anba-Mondoloni, J. (2005). Genetic characterization of the beta-glucuronidase enzyme from a human intestinal bacterium, Ruminococcus gnavus. Microbiology *151*, 2323-2330.

Bignell, G.R., Greenman, C.D., Davies, H., Butler, A.P., Edkins, S., Andrews, J.M., Buck, G., Chen, L., Beare, D., Latimer, C., et al. (2010). Signatures of mutation and selection in the cancer genome. Nature *463*, 893-898.

Boi, M., Stathis, A., Zucca, E., Inghirami, G., and Bertoni, F. (2012). Genetic alterations in systemic nodal and extranodal non-cutaneous lymphomas derived from mature T cells and natural killer cells. Cancer science *103*, 1397-1404.

Bonen, D.K., Ogura, Y., Nicolae, D.L., Inohara, N., Saab, L., Tanabe, T., Chen, F.F., Foster, S.J., Duerr, R.H., Brant, S.R., et al. (2003). Crohn's disease-associated NOD2 variants share a signaling defect in response to lipopolysaccharide and peptidoglycan. Gastroenterology *124*, 140-146.

Browning, S.R., and Browning, B.L. (2011). Haplotype phasing: existing methods and new developments. Nat Rev Genet *12*, 703-714.

Brummelkamp, T.R., Nijman, S.M.B., Dirac, A.M.G., and Bernards, R. (2003). Loss of the cylindromatosis tumour suppressor inhibits apoptosis by activating NF-[kappa]B. Nature *424*, 797-801.

Budarf, M.L., Labbe, C., David, G., and Rioux, J.D. (2009). GWA studies: rewriting the story of IBD. Trends Genet *25*, 137-146.

Buil, A., Brown, A.A., Lappalainen, T., Vinuela, A., Davies, M.N., Zheng, H.F., Richards, J.B., Glass, D., Small, K.S., Durbin, R., et al. (2014). Gene-gene and gene-environment interactions detected by transcriptome sequence analysis in twins. Nat Genet.

Cavanaugh, J. (2001). International collaboration provides convincing linkage replication in complex disease through analysis of a large pooled data set: Crohn disease and chromosome 16. Am J Hum Genet *68*, 1165-1171.

Cavanaugh, J.A., Callen, D.F., Wilson, S.R., Stanford, P.M., Sraml, M.E., Gorska, M., Crawford, J., Whitmore, S.A., Shlegel, C., Foote, S., et al. (1998). Analysis of Australian Crohn's disease pedigrees refines the localization for susceptibility to inflammatory bowel disease on chromosome 16. Ann Hum Genet *62*, 291-298.

Cleynen, I., Vazeille, E., Artieda, M., Verspaget, H.W., Szczypiorska, M., Bringer, M.A., Lakatos, P.L., Seibold, F., Parnell, K., Weersma, R.K., et al. (2014). Genetic and microbial factors modulating the ubiquitin proteasome system in inflammatory bowel disease. Gut *63*, 1265-1274.

Cobrin, G.M., and Abreu, M.T. (2005). Defects in mucosal immunity leading to Crohn's disease. Immunological reviews *206*, 277-295.

Collins, A., and Morton, N.E. (1998). Mapping a disease locus by allelic association. Proc Natl Acad Sci U S A *95*, 1741-1745.

Consortium, U.I.G., Barrett, J.C., Lee, J.C., Lees, C.W., Prescott, N.J., Anderson, C.A., Phillips, A., Wesley, E., Parnell, K., Zhang, H., et al. (2009). Genome-wide association study of ulcerative colitis identifies three new susceptibility loci, including the HNF4A region. Nat Genet *41*, 1330-1334.

Costello, C.M., Mah, N., Hasler, R., Rosenstiel, P., Waetzig, G.H., Hahn, A., Lu, T., Gurbuz, Y., Nikolaus, S., Albrecht, M., et al. (2005). Dissection of the inflammatory bowel disease transcriptome using genome-wide cDNA microarrays. PLoS Med *2*, e199.

Courtois, G., and Gilmore, T.D. (2006). Mutations in the NF-kappaB signaling pathway: implications for human disease. Oncogene *25*, 6831-6843.

De Jager, P.L., Jia, X., Wang, J., de Bakker, P.I., Ottoboni, L., Aggarwal, N.T., Piccio, L., Raychaudhuri, S., Tran, D., Aubin, C., et al. (2009). Meta-analysis of genome scans and replication identify CD6, IRF8 and TNFRSF1A as new multiple sclerosis susceptibility loci. Nat Genet *41*, 776-782.

Eckmann, L., and Karin, M. (2005). NOD2 and Crohn's disease: loss or gain of function? Immunity *22*, 661-667.

Elding, H., Lau, W., Swallow, D.M., and Maniatis, N. (2011). Dissecting the genetics of complex inheritance: linkage disequilibrium mapping provides insight into crohn disease. American journal of human genetics *89*, 798-805.

Elding, H., Lau, W., Swallow, D.M., and Maniatis, N. (2013). Refinement in localization and identification of gene regions associated with Crohn disease. Am J Hum Genet *92*, 107-113.

Ellinghaus, D., Folseraas, T., Holm, K., Ellinghaus, E., Melum, E., Balschun, T., Laerdahl, J.K., Shiryaev, A., Gotthardt, D.N., Weismuller, T.J., et al. (2013). Genome-wide association analysis in primary sclerosing cholangitis and ulcerative colitis identifies risk loci at GPR35 and TCF4. Hepatology *58*, 1074-1083.

Erlich, S., Goldshmit, Y., Lupowitz, Z., and Pinkas-Kramarski, R. (2001). ErbB-4 activation inhibits apoptosis in PC12 cells. Neuroscience *107*, 353-362.

Franchimont, D., Vermeire, S., El Housni, H., Pierik, M., Van Steen, K., Gustot, T., Quertinmont, E., Abramowicz, M., Van Gossum, A., Deviere, J., et al. (2004). Deficient host-bacteria interactions in inflammatory bowel disease? The toll-like receptor (TLR)-4 Asp299gly polymorphism is associated with Crohn's disease and ulcerative colitis. Gut *53*, 987-992.

Franke, A., Balschun, T., Karlsen, T.H., Sventoraityte, J., Nikolaus, S., Mayr, G., Domingues, F.S., Albrecht, M., Nothnagel, M., Ellinghaus, D., et al. (2008). Sequence variants in IL10, ARPC2 and multiple other loci contribute to ulcerative colitis susceptibility. Nat Genet *40*, 1319-1323.

Franke, A., Balschun, T., Sina, C., Ellinghaus, D., Hasler, R., Mayr, G., Albrecht, M., Wittig, M., Buchert, E., Nikolaus, S., et al. (2010). Genome-wide association study for ulcerative colitis identifies risk loci at 7q22 and 22q13 (IL17REL). Nat Genet *42*, 292-294.

Franke, A., McGovern, D.P., Barrett, J.C., Wang, K., Radford-Smith, G.L., Ahmad, T., Lees, C.W., Balschun, T., Lee, J., Roberts, R., et al. (2010). Genome-wide meta-analysis increases to 71 the number of confirmed Crohn's disease susceptibility loci. Nat Genet *42*, 1118-1125.

Frey, M.R., Edelblum, K.L., Mullane, M.T., Liang, D., and Polk, D.B. (2009). The ErbB4 growth factor receptor is required for colon epithelial cell survival in the presence of TNF. Gastroenterology *136*, 217-226.

Frey, M.R., Hilliard, V.C., Mullane, M.T., and Polk, D.B. (2010). ErbB4 promotes cyclooxygenase-2 expression and cell survival in colon epithelial cells. Laboratory investigation; a journal of technical methods and pathology *90*, 1415-1424.

Gassler, N., Rohr, C., Schneider, A., Kartenbeck, J., Bach, A., Obermuller, N., Otto, H.F., and Autschbach, F. (2001). Inflammatory bowel disease is associated with changes of enterocytic junctions. Am J Physiol Gastrointest Liver Physiol *281*, G216-228.

Giallourakis, C., Stoll, M., Miller, K., Hampe, J., Lander, E.S., Daly, M.J., Schreiber, S., and Rioux, J.D. (2003). IBD5 is a general risk factor for inflammatory bowel disease: replication of association with Crohn disease and identification of a novel association with ulcerative colitis. Am J Hum Genet *73*, 205-211.

Gibson, J., Tapper, W., Cox, D., Zhang, W., Pfeufer, A., Gieger, C., Wichmann, H.E., Kaab, S., Collins, A.R., Meitinger, T., et al. (2008). A multimetric approach to analysis of genome-wide association by single markers and composite likelihood. Proc Natl Acad Sci U S A *105*, 2592-2597.

Gitter, A.H., Wullstein, F., Fromm, M., and Schulzke, J.D. (2001). Epithelial barrier defects in ulcerative colitis: characterization and quantification by electrophysiological imaging. Gastroenterology *121*, 1320-1328.

Glas, J., Torok, H.P., Unterhuber, H., Radlmayr, M., and Folwaczny, C. (2003). The -295T-to-C promoter polymorphism of the IL-16 gene is associated with Crohn's disease. Clinical immunology *106*, 197-200.

Guan, Y., and Stephens, M. (2008). Practical issues in imputation-based association mapping. PLoS Genet *4*, e1000279.

Gum, J.R., Jr., Hicks, J.W., Crawley, S.C., Dahl, C.M., Yang, S.C., Roberton, A.M., and Kim, Y.S. (2003). Initiation of transcription of the MUC3A human intestinal mucin from a TATA-less promoter and comparison with the MUC3B amino terminus. J Biol Chem *278*, 49600-49609.

Hampe, J., Cuthbert, A., Croucher, P.J., Mirza, M.M., Mascheretti, S., Fisher, S., Frenzel, H., King, K., Hasselmeyer, A., MacPherson, A.J., et al. (2001). Association between insertion mutation in NOD2 gene and Crohn's disease in German and British populations. Lancet *357*, 1925-1928.

Hampe, J., Franke, A., Rosenstiel, P., Till, A., Teuber, M., Huse, K., Albrecht, M., Mayr, G., De La Vega, F.M., Briggs, J., et al. (2007). A genome-wide association

scan of nonsynonymous SNPs identifies a susceptibility variant for Crohn disease in ATG16L1. Nature genetics *39*, 207-211.

Hampe, J., Schreiber, S., Shaw, S.H., Lau, K.F., Bridger, S., Macpherson, A.J., Cardon, L.R., Sakul, H., Harris, T.J., Buckler, A., et al. (1999). A genomewide analysis provides evidence for novel linkages in inflammatory bowel disease in a large European cohort. Am J Hum Genet *64*, 808-816.

Hampe, J., Shaw, S.H., Saiz, R., Leysens, N., Lantermann, A., Mascheretti, S., Lynch, N.J., MacPherson, A.J., Bridger, S., van Deventer, S., et al. (1999). Linkage of inflammatory bowel disease to human chromosome 6p. Am J Hum Genet *65*, 1647-1655.

Hardy, R.G., Tselepis, C., Hoyland, J., Wallis, Y., Pretlow, T.P., Talbot, I., Sanders, D.S., Matthews, G., Morton, D., and Jankowski, J.A. (2002). Aberrant P-cadherin expression is an early event in hyperplastic and dysplastic transformation in the colon. Gut *50*, 513-519.

Haritunians, T., Taylor, K.D., Targan, S.R., Dubinsky, M., Ippoliti, A., Kwon, S., Guo, X., Melmed, G.Y., Berel, D., Mengesha, E., et al. (2010). Genetic predictors of medically refractory ulcerative colitis. Inflammatory bowel diseases *16*, 1830-1840.

Hill, W.G. (1974). Estimation of linkage disequilibrium in randomly mating populations. Heredity (Edinb) *33*, 229-239.

Hippisley-Cox, J., Coupland, C., Vinogradova, Y., Robson, J., May, M., and Brindle, P. (2007). Derivation and validation of QRISK, a new cardiovascular disease risk score for the United Kingdom: prospective open cohort study. BMJ *335*, 136.

Houlston, R.S., Webb, E., Broderick, P., Pittman, A.M., Di Bernardo, M.C., Lubbe, S., Chandler, I., Vijayakrishnan, J., Sullivan, K., Penegar, S., et al. (2008). Meta-analysis of genome-wide association data identifies four new susceptibility loci for colorectal cancer. Nat Genet *40*, 1426-1435.

Hugot, J.P., Chamaillard, M., Zouali, H., Lesage, S., Cezard, J.P., Belaiche, J., Almer, S., Tysk, C., O'Morain, C.A., Gassull, M., et al. (2001). Association of NOD2 leucine-rich repeat variants with susceptibility to Crohn's disease. Nature *411*, 599-603.

Hugot, J.P., Laurent-Puig, P., Gower-Rousseau, C., Olson, J.M., Lee, J.C., Beaugerie, L., Naom, I., Dupas, J.L., Van Gossum, A., Orholm, M., et al. (1996). Mapping of a susceptibility locus for Crohn's disease on chromosome 16. Nature *379*, 821-823.

Hugot, J.P., Zaccaria, I., Cavanaugh, J., Yang, H., Vermeire, S., Lappalainen, M., Schreiber, S., Annese, V., Jewell, D.P., Fowler, E.V., et al. (2007). Prevalence of CARD15/NOD2 mutations in Caucasian healthy people. Am J Gastroenterol *102*, 1259-1267.

I. Cleynen, E.V., M. Artieda, M. Szczypiorska, M. Bringer, H.W. Verspaget, P.L. Lakatos, F. Seibold, A. Tariq, R.K. Weersma, I. Arijs, S. Müller, A. Tordai, D.W.

Hommes, K. Parnell, C. Wijmenga, P. Rutgeerts, D. Lottaz, K. Van Steen, A. Darfeuille-Michaud, S. Vermeire, IBDase FP7 European Consortium (2011). Evidence for a role of the familial cylindromatosis tumor suppressor CYLD in inflammatory bowel disease. . In Presented at the 12th International Congress of Human Genetics/61st Annual Meeting of The American Society of Human Genetics. (Montreal, Canada).

Imielinski, M., Baldassano, R.N., Griffiths, A., Russell, R.K., Annese, V., Dubinsky, M., Kugathasan, S., Bradfield, J.P., Walters, T.D., Sleiman, P., et al. (2009). Common variants at five new loci associated with early-onset inflammatory bowel disease. Nature genetics *41*, 1335-1340.

Jeffreys, A.J., and Neumann, R. (2009). The rise and fall of a human recombination hot spot. Nat Genet *41*, 625-629.

Jiong, T. (2011). Genomic Discovery of recurrent CD44-SLC1A2 Gene Fusion in Gastric Cancer. In School of Medicine (Singapore: National University of Singapore), p. 180.

Johnson, R.A., Wright, K.D., Poppleton, H., Mohankumar, K.M., Finkelstein, D., Pounds, S.B., Rand, V., Leary, S.E., White, E., Eden, C., et al. (2010). Cross-species genomics matches driver mutations and cell compartments to model ependymoma. Nature *466*, 632-636.

Jones, B., and Swallow, D. (2011). The impact of cis-acting polymorphisms on the human phenotype. The HUGO Journal *5*, 13-23.

Joossens, M., Huys, G., Cnockaert, M., De Preter, V., Verbeke, K., Rutgeerts, P., Vandamme, P., and Vermeire, S. (2011). Dysbiosis of the faecal microbiota in patients with Crohn's disease and their unaffected relatives. Gut *60*, 631-637.

Jostins, L., Ripke, S., Weersma, R.K., Duerr, R.H., McGovern, D.P., Hui, K.Y., Lee, J.C., Schumm, L.P., Sharma, Y., Anderson, C.A., et al. (2012). Host-microbe interactions have shaped the genetic architecture of inflammatory bowel disease. Nature *491*, 119-124.

Kang, H.G., Jenabi, J.M., Zhang, J., Keshelava, N., Shimada, H., May, W.A., Ng, T., Reynolds, C.P., Triche, T.J., and Sorensen, P.H. (2007). E-cadherin cell-cell adhesion in ewing tumor cells mediates suppression of anoikis through activation of the ErbB4 tyrosine kinase. Cancer research *67*, 3094-3105.

Kong, A., Thorleifsson, G., Gudbjartsson, D.F., Masson, G., Sigurdsson, A., Jonasdottir, A., Walters, G.B., Jonasdottir, A., Gylfason, A., Kristinsson, K.T., et al. (2010). Fine-scale recombination rate differences between sexes, populations and individuals. Nature *467*, 1099-1103.

Kyo, K., Muto, T., Nagawa, H., Lathrop, G.M., and Nakamura, Y. (2001). Associations of distinct variants of the intestinal mucin gene MUC3A with ulcerative colitis and Crohn's disease. Journal of human genetics *46*, 5-20.

Lau, W., Kuo, T.Y., Tapper, W., Cox, S., and Collins, A. (2007). Exploiting large scale computing to construct high resolution linkage disequilibrium maps of the human genome. Bioinformatics *23*, 517-519.

Lesage, S., Zouali, H., Cezard, J.P., Colombel, J.F., Belaiche, J., Almer, S., Tysk, C., O'Morain, C., Gassull, M., Binder, V., et al. (2002). CARD15/NOD2 mutational analysis and genotype-phenotype correlation in 612 patients with inflammatory bowel disease. Am J Hum Genet *70*, 845-857.

Lessard, C.J., Adrianto, I., Ice, J.A., Wiley, G.B., Kelly, J.A., Glenn, S.B., Adler, A.J., Li, H., Rasmussen, A., Williams, A.H., et al. (2012). Identification of IRF8, TMEM39A, and IKZF3-ZPBP2 as susceptibility loci for systemic lupus erythematosus in a large-scale multiracial replication study. Am J Hum Genet *90*, 648-660.

Li, B., and Leal, S.M. (2008). Methods for detecting associations with rare variants for common diseases: application to analysis of sequence data. Am J Hum Genet *83*, 311-321.

Li, L.C., Chui, R.M., Sasaki, M., Nakajima, K., Perinchery, G., Au, H.C., Nojima, D., Carroll, P., and Dahiya, R. (2000). A single nucleotide polymorphism in the E-cadherin gene promoter alters transcriptional activities. Cancer Res *60*, 873-876.

Libioulle, C., Louis, E., Hansoul, S., Sandor, C., Farnir, F., Franchimont, D., Vermeire, S., Dewit, O., de Vos, M., Dixon, A., 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 genetics *3*, e58.

Lim, J., Thompson, J., May, R.C., Hotchin, N.A., and Caron, E. (2013). Regulator of G-Protein Signalling-14 (RGS14) Regulates the Activation of alphaMbeta2 Integrin during Phagocytosis. PLoS One *8*, e69163.

Linnebacher, M., Ostwald, C., Koczan, D., Salem, T., Schneider, B., Krohn, M., Ernst, M., and Prall, F. (2013). Single nucleotide polymorphism array analysis of microsatellite-stable, diploid/near-diploid colorectal carcinomas without the CpG island methylator phenotype. Oncology letters *5*, 173-178.

Loftus, E.V., Jr. (2004). Clinical epidemiology of inflammatory bowel disease: Incidence, prevalence, and environmental influences. Gastroenterology *126*, 1504-1517.

Maeda, S., Hsu, L.C., Liu, H., Bankston, L.A., Iimura, M., Kagnoff, M.F., Eckmann, L., and Karin, M. (2005). Nod2 mutation in Crohn's disease potentiates NF-kappaB activity and IL-1beta processing. Science *307*, 734-738.

Mai, V., and Draganov, P.V. (2009). Recent advances and remaining gaps in our knowledge of associations between gut microbiota and human health. World journal of gastroenterology : WJG *15*, 81-85.

Maniatis, N., Collins, A., and Morton, N.E. (2007). Effects of single SNPs, haplotypes, and whole-genome LD maps on accuracy of association mapping. Genet Epidemiol *31*, 179-188.

Maniatis, N., Collins, A., Xu, C.F., McCarthy, L.C., Hewett, D.R., Tapper, W., Ennis, S., Ke, X., and Morton, N.E. (2002). The first linkage disequilibrium (LD) maps: delineation of hot and cold blocks by diplotype analysis. Proc Natl Acad Sci U S A *99*, 2228-2233.

Maniatis, N., Morton, N.E., Gibson, J., Xu, C.F., Hosking, L.K., and Collins, A. (2005). The optimal measure of linkage disequilibrium reduces error in association mapping of affection status. Hum Mol Genet *14*, 145-153.

Mayer, T.Z., Simard, F.A., Cloutier, A., Vardhan, H., Dubois, C.M., and McDonald, P.P. (2013). The p38-MSK1 signaling cascade influences cytokine production through CREB and C/EBP factors in human neutrophils. J Immunol *191*, 4299-4307.

McEvoy, B.P., Montgomery, G.W., McRae, A.F., Ripatti, S., Perola, M., Spector, T.D., Cherkas, L., Ahmadi, K.R., Boomsma, D., Willemsen, G., et al. (2009). Geographical structure and differential natural selection among North European populations. Genome Res *19*, 804-814.

McGovern, D.P., Gardet, A., Torkvist, L., Goyette, P., Essers, J., Taylor, K.D., Neale, B.M., Ong, R.T., Lagace, C., Li, C., et al. (2010). Genome-wide association identifies multiple ulcerative colitis susceptibility loci. Nat Genet *42*, 332-337.

McGovern, D.P., Jones, M.R., Taylor, K.D., Marciante, K., Yan, X., Dubinsky, M., Ippoliti, A., Vasiliauskas, E., Berel, D., Derkowski, C., et al. (2010). Fucosyltransferase 2 (FUT2) non-secretor status is associated with Crohn's disease. Hum Mol Genet *19*, 3468-3476.

Milicic, A., Harrison, L.A., Goodlad, R.A., Hardy, R.G., Nicholson, A.M., Presz, M., Sieber, O., Santander, S., Pringle, J.H., Mandir, N., et al. (2008). Ectopic expression of P-cadherin correlates with promoter hypomethylation early in colorectal carcinogenesis and enhanced intestinal crypt fission in vivo. Cancer Res *68*, 7760-7768.

Montgomery, S.B., and Dermitzakis, E.T. (2011). From expression QTLs to personalized transcriptomics. Nat Rev Genet *12*, 277-282.

Morton, N., Maniatis, N., Zhang, W., Ennis, S., and Collins, A. (2007). Genome scanning by composite likelihood. Am J Hum Genet *80*, 19-28.

Morton, N.E., Zhang, W., Taillon-Miller, P., Ennis, S., Kwok, P.Y., and Collins, A. (2001). The optimal measure of allelic association. Proc Natl Acad Sci U S A *98*, 5217-5221.

Nagase, H., Visse, R., and Murphy, G. (2006). Structure and function of matrix metalloproteinases and TIMPs. Cardiovasc Res *69*, 562-573.

Nicassio, F., Corrado, N., Vissers, J.H., Areces, L.B., Bergink, S., Marteijn, J.A., Geverts, B., Houtsmuller, A.B., Vermeulen, W., Di Fiore, P.P., et al. (2007). Human USP3 is a chromatin modifier required for S phase progression and genome stability. Curr Biol *17*, 1972-1977.

Noel, A., Gutierrez-Fernandez, A., Sounni, N.E., Behrendt, N., Maquoi, E., Lund, I.K., Cal, S., Hoyer-Hansen, G., and Lopez-Otin, C. (2012). New and paradoxical roles of matrix metalloproteinases in the tumor microenvironment. Frontiers in pharmacology *3*, 140.

Ohmen, J.D., Yang, H.Y., Yamamoto, K.K., Zhao, H.Y., Ma, Y., Bentley, L.G., Huang, Z., Gerwehr, S., Pressman, S., McElree, C., et al. (1996). Susceptibility locus for inflammatory bowel disease on chromosome 16 has a role in Crohn's disease, but not in ulcerative colitis. Hum Mol Genet *5*, 1679-1683.

Olden, K. (2006). Toxicogenomics--a new systems toxicology approach to understanding of gene-environment interactions. Ann N Y Acad Sci *1076*, 703-706.

Orholm, M., Munkholm, P., Langholz, E., Nielsen, O.H., Sorensen, T.I., and Binder, V. (1991). Familial occurrence of inflammatory bowel disease. N Engl J Med *324*, 84-88.

Ott, S.J., Musfeldt, M., Wenderoth, D.F., Hampe, J., Brant, O., Folsch, U.R., Timmis, K.N., and Schreiber, S. (2004). Reduction in diversity of the colonic mucosa associated bacterial microflora in patients with active inflammatory bowel disease. Gut *53*, 685-693.

Ouburg, S., Mallant-Hent, R., Crusius, J.B., van Bodegraven, A.A., Mulder, C.J., Linskens, R., Pena, A.S., and Morre, S.A. (2005). The toll-like receptor 4 (TLR4) Asp299Gly polymorphism is associated with colonic localisation of Crohn's disease without a major role for the Saccharomyces cerevisiae mannan-LBP-CD14-TLR4 pathway. Gut *54*, 439-440.

Parkes, M., Barmada, M.M., Satsangi, J., Weeks, D.E., Jewell, D.P., and Duerr, R.H. (2000). The IBD2 locus shows linkage heterogeneity between ulcerative colitis and Crohn disease. Am J Hum Genet *67*, 1605-1610.

Pennacchio, L.A., Ahituv, N., Moses, A.M., Prabhakar, S., Nobrega, M.A., Shoukry, M., Minovitsky, S., Dubchak, I., Holt, A., Lewis, K.D., et al. (2006). In vivo enhancer analysis of human conserved non-coding sequences. Nature *444*, 499-502.

Politopoulos, I., Gibson, J., Tapper, W., Ennis, S., Eccles, D., and Collins, A. (2011). Genome-wide association of breast cancer: composite likelihood with imputed genotypes. Eur J Hum Genet *19*, 194-199.

Porter, T.R., Richards, F.M., Houlston, R.S., Evans, D.G., Jankowski, J.A., Macdonald, F., Norbury, G., Payne, S.J., Fisher, S.A., Tomlinson, I., et al. (2002). Contribution of cyclin d1 (CCND1) and E-cadherin (CDH1) polymorphisms to familial and sporadic colorectal cancer. Oncogene *21*, 1928-1933.

Pratt, W.S., Crawley, S., Hicks, J., Ho, J., Nash, M., Kim, Y.S., Gum, J.R., and Swallow, D.M. (2000). Multiple transcripts of MUC3: evidence for two genes, MUC3A and MUC3B. Biochemical and biophysical research communications *275*, 916-923.

Ramos, R.G., and Olden, K. (2008). Gene-environment interactions in the development of complex disease phenotypes. International journal of environmental research and public health *5*, 4-11.

Reiley, W.W., Jin, W., Lee, A.J., Wright, A., Wu, X., Tewalt, E.F., Leonard, T.O., Norbury, C.C., Fitzpatrick, L., Zhang, M., et al. (2007). Deubiquitinating enzyme

CYLD negatively regulates the ubiquitin-dependent kinase Tak1 and prevents abnormal T cell responses. J Exp Med *204*, 1475-1485.

Renner, M., Bergmann, G., Krebs, I., End, C., Lyer, S., Hilberg, F., Helmke, B., Gassler, N., Autschbach, F., Bikker, F., et al. (2007). DMBT1 confers mucosal protection in vivo and a deletion variant is associated with Crohn's disease. Gastroenterology *133*, 1499-1509.

Rioux, J.D., Xavier, R.J., Taylor, K.D., Silverberg, M.S., Goyette, P., Huett, A., Green, T., Kuballa, P., Barmada, M.M., Datta, L.W., 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 Merikangas, K. (1996). The future of genetic studies of complex human diseases. Science *273*, 1516-1517.

Risques, R.A., Lai, L.A., Brentnall, T.A., Li, L., Feng, Z., Gallaher, J., Mandelson, M.T., Potter, J.D., Bronner, M.P., and Rabinovitch, P.S. (2008). Ulcerative colitis is a disease of accelerated colon aging: evidence from telomere attrition and DNA damage. Gastroenterology *135*, 410-418.

Rivas, M.A., Beaudoin, M., Gardet, A., Stevens, C., Sharma, Y., Zhang, C.K., Boucher, G., Ripke, S., Ellinghaus, D., Burtt, N., et al. (2011). Deep resequencing of GWAS loci identifies independent rare variants associated with inflammatory bowel disease. Nature genetics *43*, 1066-1073.

Sartor, R.B. (2006). Mechanisms of disease: pathogenesis of Crohn's disease and ulcerative colitis. Nature clinical practice. Gastroenterology & hepatology *3*, 390-407.

Satsangi, J., Parkes, M., Louis, E., Hashimoto, L., Kato, N., Welsh, K., Terwilliger, J.D., Lathrop, G.M., Bell, J.I., and Jewell, D.P. (1996). Two stage genome-wide search in inflammatory bowel disease provides evidence for susceptibility loci on chromosomes 3, 7 and 12. Nat Genet *14*, 199-202.

Satsangi, J.R., William M.C. *; Jewell, Derek P. (1994). The prevalence of inflammatory bowel disease in relatives of patients with Crohn’s disease. European Journal of Gastroenterology and Hepatology *6*, 413-416.

Schmitz, H., Barmeyer, C., Fromm, M., Runkel, N., Foss, H.D., Bentzel, C.J., Riecken, E.O., and Schulzke, J.D. (1999). Altered tight junction structure contributes to the impaired epithelial barrier function in ulcerative colitis. Gastroenterology *116*, 301-309.

Schneider, M.R., Dahlhoff, M., Horst, D., Hirschi, B., Trulzsch, K., Muller-Hocker, J., Vogelmann, R., Allgauer, M., Gerhard, M., Steininger, S., et al. (2010). A Key Role for E-cadherin in Intestinal Homeostasis and Paneth Cell Maturation. PLoS One *5*, e14325.

Schreiber, S., Rosenstiel, P., Albrecht, M., Hampe, J., and Krawczak, M. (2005). Genetics of Crohn disease, an archetypal inflammatory barrier disease. Nat Rev Genet *6*, 376-388.

Schurr, E., and Gros, P. (2009). A common genetic fingerprint in leprosy and Crohn's disease? N Engl J Med *361*, 2666-2668.

Shih Ie, M., Nakayama, K., Wu, G., Nakayama, N., Zhang, J., and Wang, T.L. (2011). Amplification of the ch19p13.2 NACC1 locus in ovarian high-grade serous carcinoma. Modern pathology : an official journal of the United States and Canadian Academy of Pathology, Inc *24*, 638-645.

Silverberg, M.S., Cho, J.H., Rioux, J.D., McGovern, D.P., Wu, J., Annese, V., Achkar, J.P., Goyette, P., Scott, R., Xu, W., et al. (2009). Ulcerative colitis-risk loci on chromosomes 1p36 and 12q15 found by genome-wide association study. Nat Genet *41*, 216-220.

Silverman, E.K., and Palmer, L.J. (2000). Case-control association studies for the genetics of complex respiratory diseases. Am J Respir Cell Mol Biol *22*, 645-648.

Slatkin, M. (2008). Linkage disequilibrium--understanding the evolutionary past and mapping the medical future. Nat Rev Genet *9*, 477-485.

Starr, A., Greif, J., Vexler, A., Ashkenazy-Voghera, M., Gladesh, V., Rubin, C., Kerber, G., Marmor, S., Lev-Ari, S., Inbar, M., et al. (2006). ErbB4 increases the proliferation potential of human lung cancer cells and its blockage can be used as a target for anti-cancer therapy. International journal of cancer. Journal international du cancer *119*, 269-274.

Steinberg, M.W., Turovskaya, O., Shaikh, R.B., Kim, G., McCole, D.F., Pfeffer, K., Murphy, K.M., Ware, C.F., and Kronenberg, M. (2008). A crucial role for HVEM and BTLA in preventing intestinal inflammation. The Journal of experimental medicine *205*, 1463-1476.

Sterne, J.A., White, I.R., Carlin, J.B., Spratt, M., Royston, P., Kenward, M.G., Wood, A.M., and Carpenter, J.R. (2009). Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. BMJ *338*, b2393.

Tapper, W. (2007). Linkage disequilibrium maps and location databases. Methods Mol Biol *376*, 23-45.

Tapper, W., Collins, A., Gibson, J., Maniatis, N., Ennis, S., and Morton, N.E. (2005). A map of the human genome in linkage disequilibrium units. Proc Natl Acad Sci U S A *102*, 11835-11839.

Targan, S.R., and Karp, L.C. (2005). Defects in mucosal immunity leading to ulcerative colitis. Immunological reviews *206*, 296-305.

Thompson, A.I., and Lees, C.W. (2011). Genetics of ulcerative colitis. Inflammatory bowel diseases *17*, 831-848.

Tysk, C., Lindberg, E., Jarnerot, G., and Floderus-Myrhed, B. (1988). 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 *29*, 990-996.

van Heel, D.A., Dechairo, B.M., Dawson, G., McGovern, D.P., Negoro, K., Carey, A.H., Cardon, L.R., Mackay, I., Jewell, D.P., and Lench, N.J. (2003). The IBD6

Crohn's disease locus demonstrates complex interactions with CARD15 and IBD5 disease-associated variants. Hum Mol Genet *12*, 2569-2575.

VanLiere, J.M., and Rosenberg, N.A. (2008). Mathematical properties of the r2 measure of linkage disequilibrium. Theor Popul Biol *74*, 130-137.

Villani, A.C., Lemire, M., Thabane, M., Belisle, A., Geneau, G., Garg, A.X., Clark, W.F., Moayyedi, P., Collins, S.M., Franchimont, D., et al. (2010). Genetic risk factors for post-infectious irritable bowel syndrome following a waterborne outbreak of gastroenteritis. Gastroenterology *138*, 1502-1513.

Wang, K., Yuen, S.T., Xu, J., Lee, S.P., Yan, H.H., Shi, S.T., Siu, H.C., Deng, S., Chu, K.M., Law, S., et al. (2014). Whole-genome sequencing and comprehensive molecular profiling identify new driver mutations in gastric cancer. Nat Genet *46*, 573-582.

Webb, A.J., Berg, I.L., and Jeffreys, A. (2008). Sperm cross-over activity in regions of the human genome showing extreme breakdown of marker association. Proc Natl Acad Sci U S A *105*, 10471-10476.

Wehkamp, J., Schmid, M., and Stange, E.F. (2007). Defensins and other antimicrobial peptides in inflammatory bowel disease. Current opinion in gastroenterology *23*, 370-378.

Weiss, B., Lebowitz, O., Fidder, H.H., Maza, I., Levine, A., Shaoul, R., Reif, S., Bujanover, Y., and Karban, A. (2010). Response to medical treatment in patients with Crohn's disease: the role of NOD2/CRAD15, disease phenotype, and age of diagnosis. Dig Dis Sci *55*, 1674-1680.

Wheeler, J.M., Kim, H.C., Efstathiou, J.A., Ilyas, M., Mortensen, N.J., and Bodmer, W.F. (2001). Hypermethylation of the promoter region of the E-cadherin gene (CDH1) in sporadic and ulcerative colitis associated colorectal cancer. Gut *48*, 367-371.

Winney, B., Boumertit, A., Day, T., Davison, D., Echeta, C., Evseeva, I., Hutnik, K., Leslie, S., Nicodemus, K., Royrvik, E.C., et al. (2012). People of the British Isles: preliminary analysis of genotypes and surnames in a UK-control population. Eur J Hum Genet *20*, 203-210.

WTCCC (2007). Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature *447*, 661-678.

Xavier, R.J., and Podolsky, D.K. (2007). Unravelling the pathogenesis of inflammatory bowel disease. Nature *448*, 427-434.

Yang, S.K., Hong, M., Zhao, W., Jung, Y., Tayebi, N., Ye, B.D., Kim, K.J., Park, S.H., Lee, I., Shin, H.D., et al. (2013). Genome-wide association study of ulcerative colitis in Koreans suggests extensive overlapping of genetic susceptibility with Caucasians. Inflammatory bowel diseases *19*, 954-966.

Zervoudi, E., Saridakis, E., Birtley, J.R., Seregin, S.S., Reeves, E., Kokkala, P., Aldhamen, Y.A., Amalfitano, A., Mavridis, I.M., James, E., et al. (2013). Rationally designed inhibitor targeting antigen-trimming aminopeptidases enhances antigen

presentation and cytotoxic T-cell responses. Proc Natl Acad Sci U S A *110*, 19890-19895.

Zhang, F., Liu, H., Chen, S., Low, H., Sun, L., Cui, Y., Chu, T., Li, Y., Fu, X., Yu, Y., et al. (2011). Identification of two new loci at IL23R and RAB32 that influence susceptibility to leprosy. Nat Genet *43*, 1247-1251.

Zhang, F.R., Huang, W., Chen, S.M., Sun, L.D., Liu, H., Li, Y., Cui, Y., Yan, X.X., Yang, H.T., Yang, R.D., et al. (2009). Genomewide association study of leprosy. N Engl J Med *361*, 2609-2618.

Zhang, J., Stirling, B., Temmerman, S.T., Ma, C.A., Fuss, I.J., Derry, J.M., and Jain, A. (2006). Impaired regulation of NF-kappaB and increased susceptibility to colitis-associated tumorigenesis in CYLD-deficient mice. J Clin Invest *116*, 3042-3049.

Zhao, B., Takami, M., Yamada, A., Wang, X., Koga, T., Hu, X., Tamura, T., Ozato, K., Choi, Y., Ivashkiv, L.B., et al. (2009). Interferon regulatory factor-8 regulates bone metabolism by suppressing osteoclastogenesis. Nat Med *15*, 1066-1071.

Zheng, X., Levine, D., Shen, J., Gogarten, S.M., Laurie, C., and Weir, B.S. (2012). A high-performance computing toolset for relatedness and principal component analysis of SNP data. Bioinformatics *28*, 3326-3328.

Zhernakova, A., Festen, E.M., Franke, L., Trynka, G., van Diemen, C.C., Monsuur, A.J., Bevova, M., Nijmeijer, R.M., van 't Slot, R., Heijmans, R., et al. (2008). Genetic analysis of innate immunity in Crohn's disease and ulcerative colitis identifies two susceptibility loci harboring CARD9 and IL18RAP. American journal of human genetics *82*, 1202-1210.

Appendix I- Different LD metrics

Formula for Covariance *D*:

$$D = f_{11}f_{22} - f_{12}f_{21}$$

where *f* is the frequency of the inferred haplotype.

Formula for *D'*:

When D is negative, $D' = \frac{|D|}{\min[QR, (1-Q)(1-R)]}$

When D is positive, $D' = \frac{|D|}{\min[Q(1-R), R(1-Q)]}$

Formula for association *ρ*:

$$\rho = \frac{D}{Q(1-R)}$$

where *D* is the covariance and *Q* and *R* are the allele frequencies.

Formula for correlation *r*:

$$r = \frac{D}{\sqrt{Q(1-Q)R(1-R)}}$$

where *D* is the covariance and *Q* and *R* are the allele frequencies.

Formula for correlation coefficient r^2:

$$r^2 = \frac{D^2}{Q(1-Q)R(1-R)}$$

where D is the covariance and Q and R are the allele frequencies.

Appendix II- BioMart Analysis from the Crohn's Disease GWAS described in Chapter III.

Ensembl Gene ID	GO Term Accession	GO Term Name	Gene Symbol
ENSG00000077522	GO:0042981	regulation of apoptotic process	*ACTN2*
ENSG00000077522	GO:0007155	cell adhesion	*ACTN2*
ENSG00000077522	GO:0006936	muscle contraction	*ACTN2*
ENSG00000077522	GO:0007268	synaptic transmission	*ACTN2*
ENSG00000077522	GO:0048041	focal adhesion assembly	*ACTN2*
ENSG00000077522	GO:0030168	platelet activation	*ACTN2*
ENSG00000077522	GO:0051289	protein homotetramerization	*ACTN2*
ENSG00000077522	GO:0007596	blood coagulation	*ACTN2*
ENSG00000077522	GO:0002576	platelet degranulation	*ACTN2*
ENSG00000077522	GO:0030035	microspike assembly	*ACTN2*
ENSG00000077522	GO:0030049	muscle filament sliding	*ACTN2*
ENSG00000077522	GO:0005737	cytoplasm	*ACTN2*
ENSG00000077522	GO:0005576	extracellular region	*ACTN2*
ENSG00000077522	GO:0005829	cytosol	*ACTN2*
ENSG00000077522	GO:0005856	cytoskeleton	*ACTN2*
ENSG00000077522	GO:0030018	Z disc	*ACTN2*
ENSG00000077522	GO:0043197	dendritic spine	*ACTN2*
ENSG00000077522	GO:0030175	filopodium	*ACTN2*
ENSG00000077522	GO:0005925	focal adhesion	*ACTN2*
ENSG00000077522	GO:0031143	pseudopodium	*ACTN2*
ENSG00000077522	GO:0030017	sarcomere	*ACTN2*
ENSG00000077522	GO:0005884	actin filament	*ACTN2*
ENSG00000077522	GO:0030864	cortical actin cytoskeleton	*ACTN2*
ENSG00000077522	GO:0031093	platelet alpha granule lumen	*ACTN2*
ENSG00000077522	GO:0005515	protein binding	*ACTN2*
ENSG00000077522	GO:0005509	calcium ion binding	*ACTN2*
ENSG00000077522	GO:0005178	integrin binding	*ACTN2*
ENSG00000077522	GO:0046983	protein dimerization activity	*ACTN2*
ENSG00000077522	GO:0051015	actin filament binding	*ACTN2*
ENSG00000077522	GO:0042802	identical protein binding	*ACTN2*
ENSG00000077522	GO:0030375	thyroid hormone receptor coactivator activity	*ACTN2*
ENSG00000077522	GO:0044325	ion channel binding	*ACTN2*
ENSG00000077522	GO:0008307	structural constituent of muscle	*ACTN2*
ENSG00000077522	GO:0031432	titin binding	*ACTN2*
ENSG00000077522	GO:0070080	titin Z domain binding	*ACTN2*
ENSG00000077522	GO:0051370	ZASP binding	*ACTN2*
ENSG00000077522	GO:0051374	FATZ 1 binding	*ACTN2*
ENSG00000077522	GO:0030374	ligand-dependent nuclear receptor transcription coactivator activity	*ACTN2*
ENSG00000077522	GO:0003779	actin binding	*ACTN2*
ENSG00000138031	GO:0055085	transmembrane transport	*ADCY3*
ENSG00000138031	GO:0007165	signal transduction	*ADCY3*
ENSG00000138031	GO:0044281	small molecule metabolic process	*ADCY3*
ENSG00000138031	GO:0035556	intracellular signal transduction	*ADCY3*
ENSG00000138031	GO:0007143	female meiosis	*ADCY3*
ENSG00000138031	GO:0050896	response to stimulus	*ADCY3*
ENSG00000138031	GO:0007189	activation of adenylate cyclase activity by G-protein signaling pathway	*ADCY3*
ENSG00000138031	GO:0006833	water transport	*ADCY3*

ENSG00000138031	GO:0007268	synaptic transmission	*ADCY3*
ENSG00000138031	GO:0006184	GTP catabolic process	*ADCY3*
ENSG00000138031	GO:0009190	cyclic nucleotide biosynthetic process	*ADCY3*
ENSG00000138031	GO:0007608	sensory perception of smell	*ADCY3*
ENSG00000138031	GO:0007193	inhibition of adenylate cyclase activity by G-protein signaling pathway	*ADCY3*
ENSG00000138031	GO:0007173	epidermal growth factor receptor signaling pathway	*ADCY3*
ENSG00000138031	GO:0071377	cellular response to glucagon stimulus	*ADCY3*
ENSG00000138031	GO:0007202	activation of phospholipase C activity	*ADCY3*
ENSG00000138031	GO:0008543	fibroblast growth factor receptor signaling pathway	*ADCY3*
ENSG00000138031	GO:0048011	nerve growth factor receptor signaling pathway	*ADCY3*
ENSG00000138031	GO:0006112	energy reserve metabolic process	*ADCY3*
ENSG00000138031	GO:0006171	cAMP biosynthetic process	*ADCY3*
ENSG00000138031	GO:0034199	activation of protein kinase A activity	*ADCY3*
ENSG00000138031	GO:0005634	nucleus	*ADCY3*
ENSG00000138031	GO:0005737	cytoplasm	*ADCY3*
ENSG00000138031	GO:0005886	plasma membrane	*ADCY3*
ENSG00000138031	GO:0005624	membrane fraction	*ADCY3*
ENSG00000138031	GO:0005887	integral to plasma membrane	*ADCY3*
ENSG00000138031	GO:0045121	membrane raft	*ADCY3*
ENSG00000138031	GO:0000166	nucleotide binding	*ADCY3*
ENSG00000138031	GO:0005524	ATP binding	*ADCY3*
ENSG00000138031	GO:0046872	metal ion binding	*ADCY3*
ENSG00000138031	GO:0005516	calmodulin binding	*ADCY3*
ENSG00000138031	GO:0003924	GTPase activity	*ADCY3*
ENSG00000138031	GO:0016849	phosphorus-oxygen lyase activity	*ADCY3*
ENSG00000138031	GO:0004016	adenylate cyclase activity	*ADCY3*
ENSG00000138031	GO:0008294	calcium- and calmodulin-responsive adenylate cyclase activity	*ADCY3*
ENSG00000138031	GO:0005515	protein binding	*ADCY3*
ENSG00000144891	GO:0050729	positive regulation of inflammatory response	*AGTR1*
ENSG00000144891	GO:0051482	elevation of cytosolic calcium ion concentration involved in G-protein signaling coupled to IP3 second messenger	*AGTR1*
ENSG00000144891	GO:0007186	G-protein coupled receptor signaling pathway	*AGTR1*
ENSG00000144891	GO:0001558	regulation of cell growth	*AGTR1*
ENSG00000144891	GO:0007204	elevation of cytosolic calcium ion concentration	*AGTR1*
ENSG00000144891	GO:0019722	calcium-mediated signaling	*AGTR1*
ENSG00000144891	GO:0042127	regulation of cell proliferation	*AGTR1*
ENSG00000144891	GO:0007266	Rho protein signal transduction	*AGTR1*
ENSG00000144891	GO:2000379	positive regulation of reactive oxygen species metabolic process	*AGTR1*
ENSG00000144891	GO:0007200	activation of phospholipase C activity by G-protein coupled receptor protein signaling pathway coupled to IP3 second messenger	*AGTR1*
ENSG00000144891	GO:0001822	kidney development	*AGTR1*
ENSG00000144891	GO:0060326	cell chemotaxis	*AGTR1*
ENSG00000144891	GO:0010744	positive regulation of macrophage derived foam cell differentiation	*AGTR1*
ENSG00000144891	GO:0032270	positive regulation of cellular protein metabolic process	*AGTR1*
ENSG00000144891	GO:0019229	regulation of vasoconstriction	*AGTR1*

ENSG00000144891	GO:0050727	regulation of inflammatory response	*AGTR1*
ENSG00000144891	GO:0002034	regulation of blood vessel size by renin-angiotensin	*AGTR1*
ENSG00000144891	GO:0010873	positive regulation of cholesterol esterification	*AGTR1*
ENSG00000144891	GO:0034374	low-density lipoprotein particle remodeling	*AGTR1*
ENSG00000144891	GO:0002018	renin-angiotensin regulation of aldosterone production	*AGTR1*
ENSG00000144891	GO:0033864	positive regulation of NAD(P)H oxidase activity	*AGTR1*
ENSG00000144891	GO:0035813	regulation of renal sodium excretion	*AGTR1*
ENSG00000144891	GO:0003081	regulation of systemic arterial blood pressure by renin-angiotensin	*AGTR1*
ENSG00000144891	GO:0042312	regulation of vasodilation	*AGTR1*
ENSG00000144891	GO:0032430	positive regulation of phospholipase A2 activity	*AGTR1*
ENSG00000144891	GO:0016021	integral to membrane	*AGTR1*
ENSG00000144891	GO:0005886	plasma membrane	*AGTR1*
ENSG00000144891	GO:0005887	integral to plasma membrane	*AGTR1*
ENSG00000144891	GO:0004872	receptor activity	*AGTR1*
ENSG00000144891	GO:0005515	protein binding	*AGTR1*
ENSG00000144891	GO:0046982	protein heterodimerization activity	*AGTR1*
ENSG00000144891	GO:0004871	signal transducer activity	*AGTR1*
ENSG00000144891	GO:0004930	G-protein coupled receptor activity	*AGTR1*
ENSG00000144891	GO:0010698	acetyltransferase activator activity	*AGTR1*
ENSG00000144891	GO:0004945	angiotensin type II receptor activity	*AGTR1*
ENSG00000144891	GO:0031711	bradykinin receptor binding	*AGTR1*
ENSG00000144891	GO:0001596	angiotensin type I receptor activity	*AGTR1*
ENSG00000151572	GO:0006811	ion transport	*ANO4*
ENSG00000151572	GO:0016020	membrane	*ANO4*
ENSG00000151572	GO:0016021	integral to membrane	*ANO4*
ENSG00000151572	GO:0034707	chloride channel complex	*ANO4*
ENSG00000151572	GO:0005216	ion channel activity	*ANO4*
ENSG00000151572	GO:0005254	chloride channel activity	*ANO4*
ENSG00000262139	GO:0006811	ion transport	*ANO4*
ENSG00000262139	GO:0016020	membrane	*ANO4*
ENSG00000262139	GO:0016021	integral to membrane	*ANO4*
ENSG00000262139	GO:0034707	chloride channel complex	*ANO4*
ENSG00000262139	GO:0005216	ion channel activity	*ANO4*
ENSG00000262139	GO:0005254	chloride channel activity	*ANO4*
ENSG00000150347	GO:0060612	adipose tissue development	*ARID5B*
ENSG00000150347	GO:0035264	multicellular organism growth	*ARID5B*
ENSG00000150347	GO:0045892	negative regulation of transcription, DNA-dependent	*ARID5B*
ENSG00000150347	GO:0009791	post-embryonic development	*ARID5B*
ENSG00000150347	GO:0008584	male gonad development	*ARID5B*
ENSG00000150347	GO:0008585	female gonad development	*ARID5B*
ENSG00000150347	GO:0006807	nitrogen compound metabolic process	*ARID5B*
ENSG00000150347	GO:0048705	skeletal system morphogenesis	*ARID5B*
ENSG00000150347	GO:0010761	fibroblast migration	*ARID5B*
ENSG00000150347	GO:0001889	liver development	*ARID5B*
ENSG00000150347	GO:0048008	platelet-derived growth factor receptor signaling pathway	*ARID5B*
ENSG00000150347	GO:0045444	fat cell differentiation	*ARID5B*
ENSG00000150347	GO:0030325	adrenal gland development	*ARID5B*
ENSG00000150347	GO:0048468	cell development	*ARID5B*

ENSG00000150347	GO:0001822	kidney development	*ARID5B*
ENSG00000150347	GO:0060021	palate development	*ARID5B*
ENSG00000150347	GO:0060325	face morphogenesis	*ARID5B*
ENSG00000150347	GO:0048644	muscle organ morphogenesis	*ARID5B*
ENSG00000150347	GO:0051091	positive regulation of sequence-specific DNA binding transcription factor activity	*ARID5B*
ENSG00000150347	GO:0060613	fat pad development	*ARID5B*
ENSG00000150347	GO:0005634	nucleus	*ARID5B*
ENSG00000150347	GO:0005622	intracellular	*ARID5B*
ENSG00000150347	GO:0003677	DNA binding	*ARID5B*
ENSG00000150347	GO:0005515	protein binding	*ARID5B*
ENSG00000150347	GO:0044212	transcription regulatory region DNA binding	*ARID5B*
ENSG00000150347	GO:0003713	transcription coactivator activity	*ARID5B*
ENSG00000172995	GO:0034605	cellular response to heat	*ARPP21*
ENSG00000172995	GO:0003676	nucleic acid binding	*ARPP21*
ENSG00000172995	GO:0008150	biological_process	*ARPP21*
ENSG00000172995	GO:0005575	cellular_component	*ARPP21*
ENSG00000172995	GO:0005737	cytoplasm	*ARPP21*
ENSG00000172995	GO:0003674	molecular_function	*ARPP21*
ENSG00000172995	GO:0005516	calmodulin binding	*ARPP21*
ENSG00000104043	GO:0006812	cation transport	*ATP8B4*
ENSG00000104043	GO:0015914	phospholipid transport	*ATP8B4*
ENSG00000104043	GO:0016020	membrane	*ATP8B4*
ENSG00000104043	GO:0016021	integral to membrane	*ATP8B4*
ENSG00000104043	GO:0000166	nucleotide binding	*ATP8B4*
ENSG00000104043	GO:0005524	ATP binding	*ATP8B4*
ENSG00000104043	GO:0016787	hydrolase activity	*ATP8B4*
ENSG00000104043	GO:0046872	metal ion binding	*ATP8B4*
ENSG00000104043	GO:0000287	magnesium ion binding	*ATP8B4*
ENSG00000104043	GO:0016820	hydrolase activity, acting on acid anhydrides, catalyzing transmembrane movement of substances	*ATP8B4*
ENSG00000104043	GO:0015662	ATPase activity, coupled to transmembrane movement of ions, phosphorylative mechanism	*ATP8B4*
ENSG00000104043	GO:0004012	phospholipid-translocating ATPase activity	*ATP8B4*
ENSG00000176697	GO:0007411	axon guidance	*BDNF*
ENSG00000176697	GO:0007399	nervous system development	*BDNF*
ENSG00000176697	GO:0045666	positive regulation of neuron differentiation	*BDNF*
ENSG00000176697	GO:0048839	inner ear development	*BDNF*
ENSG00000176697	GO:0007631	feeding behavior	*BDNF*
ENSG00000176697	GO:0042493	response to drug	*BDNF*
ENSG00000176697	GO:0007412	axon target recognition	*BDNF*
ENSG00000176697	GO:0014047	glutamate secretion	*BDNF*
ENSG00000176697	GO:0048167	regulation of synaptic plasticity	*BDNF*
ENSG00000176697	GO:0001657	ureteric bud development	*BDNF*
ENSG00000176697	GO:0016358	dendrite development	*BDNF*
ENSG00000176697	GO:0043524	negative regulation of neuron apoptosis	*BDNF*
ENSG00000176697	GO:0006916	anti-apoptosis	*BDNF*
ENSG00000176697	GO:0007611	learning or memory	*BDNF*
ENSG00000176697	GO:0021675	nerve development	*BDNF*
ENSG00000176697	GO:0019222	regulation of metabolic process	*BDNF*
ENSG00000176697	GO:0007406	negative regulation of neuroblast proliferation	*BDNF*
ENSG00000176697	GO:0042596	fear response	*BDNF*
ENSG00000176697	GO:0046668	regulation of retinal cell programmed cell death	*BDNF*

ENSG00000176697	GO:0008038	neuron recognition	*BDNF*
ENSG00000176697	GO:0005576	extracellular region	*BDNF*
ENSG00000176697	GO:0016023	cytoplasmic membrane-bounded vesicle	*BDNF*
ENSG00000176697	GO:0005102	receptor binding	*BDNF*
ENSG00000176697	GO:0008083	growth factor activity	*BDNF*
ENSG00000176697	GO:0043523	regulation of neuron apoptosis	*BDNF*
ENSG00000176697	GO:0007610	behavior	*BDNF*
ENSG00000176697	GO:0005515	protein binding	*BDNF*
ENSG00000178409			*BEND3*
ENSG00000166164	GO:0007049	cell cycle	*BRD7*
ENSG00000166164	GO:0016055	Wnt receptor signaling pathway	*BRD7*
ENSG00000166164	GO:0006357	regulation of transcription from RNA polymerase II promoter	*BRD7*
ENSG00000166164	GO:0008285	negative regulation of cell proliferation	*BRD7*
ENSG00000166164	GO:0045893	positive regulation of transcription, DNA-dependent	*BRD7*
ENSG00000166164	GO:2000134	negative regulation of G1/S transition of mitotic cell cycle	*BRD7*
ENSG00000166164	GO:0045892	negative regulation of transcription, DNA-dependent	*BRD7*
ENSG00000166164	GO:0035066	positive regulation of histone acetylation	*BRD7*
ENSG00000166164	GO:0005634	nucleus	*BRD7*
ENSG00000166164	GO:0005737	cytoplasm	*BRD7*
ENSG00000166164	GO:0005515	protein binding	*BRD7*
ENSG00000166164	GO:0003714	transcription corepressor activity	*BRD7*
ENSG00000166164	GO:0042393	histone binding	*BRD7*
ENSG00000166164	GO:0008134	transcription factor binding	*BRD7*
ENSG00000166164	GO:0044212	transcription regulatory region DNA binding	*BRD7*
ENSG00000166164	GO:0002039	p53 binding	*BRD7*
ENSG00000166164	GO:0003713	transcription coactivator activity	*BRD7*
ENSG00000166164	GO:0070577	histone acetyl-lysine binding	*BRD7*
ENSG00000186265	GO:0030889	negative regulation of B cell proliferation	*BTLA*
ENSG00000186265	GO:0046642	negative regulation of alpha-beta T cell proliferation	*BTLA*
ENSG00000186265	GO:0031295	T cell costimulation	*BTLA*
ENSG00000186265	GO:0002768	immune response-regulating cell surface receptor signaling pathway	*BTLA*
ENSG00000186265	GO:0042130	negative regulation of T cell proliferation	*BTLA*
ENSG00000186265	GO:0007166	cell surface receptor signaling pathway	*BTLA*
ENSG00000186265	GO:0005886	plasma membrane	*BTLA*
ENSG00000186265	GO:0005887	integral to plasma membrane	*BTLA*
ENSG00000186265	GO:0009897	external side of plasma membrane	*BTLA*
ENSG00000186265	GO:0004872	receptor activity	*BTLA*
ENSG00000186265	GO:0005515	protein binding	*BTLA*
ENSG00000183346			*C10orf107*
ENSG00000214215			*C12orf74*
ENSG00000204669	GO:0016020	membrane	*C9orf57*
ENSG00000204669	GO:0016021	integral to membrane	*C9orf57*
ENSG00000155621			*C9orf85*
ENSG00000157445	GO:0006811	ion transport	*CACNA2D3*
ENSG00000157445	GO:0006816	calcium ion transport	*CACNA2D3*
ENSG00000157445	GO:0051925	regulation of calcium ion transport via voltage-gated calcium channel activity	*CACNA2D3*
ENSG00000157445	GO:0016020	membrane	*CACNA2D3*
ENSG00000157445	GO:0016021	integral to membrane	*CACNA2D3*
ENSG00000157445	GO:0005515	protein binding	*CACNA2D3*

ENSG00000157445	GO:0046872	metal ion binding	*CACNA2D3*
ENSG00000157445	GO:0005245	voltage-gated calcium channel activity	*CACNA2D3*
ENSG00000157445	GO:0005244	voltage-gated ion channel activity	*CACNA2D3*
ENSG00000157445	GO:0005262	calcium channel activity	*CACNA2D3*
ENSG00000111530	GO:0016567	protein ubiquitination	*CAND1*
ENSG00000111530	GO:0030154	cell differentiation	*CAND1*
ENSG00000111530	GO:0006355	regulation of transcription, DNA-dependent	*CAND1*
ENSG00000111530	GO:0043086	negative regulation of catalytic activity	*CAND1*
ENSG00000111530	GO:0045899	positive regulation of RNA polymerase II transcriptional preinitiation complex assembly	*CAND1*
ENSG00000111530	GO:0005634	nucleus	*CAND1*
ENSG00000111530	GO:0000151	ubiquitin ligase complex	*CAND1*
ENSG00000111530	GO:0005488	binding	*CAND1*
ENSG00000111530	GO:0005515	protein binding	*CAND1*
ENSG00000111530	GO:0017025	TBP-class protein binding	*CAND1*
ENSG00000123106	GO:0005634	nucleus	*CCDC91*
ENSG00000123106	GO:0005794	Golgi apparatus	*CCDC91*
ENSG00000123106	GO:0042802	identical protein binding	*CCDC91*
ENSG00000123106	GO:0015031	protein transport	*CCDC91*
ENSG00000123106	GO:0016020	membrane	*CCDC91*
ENSG00000101017	GO:0050776	regulation of immune response	*CD40*
ENSG00000101017	GO:0007165	signal transduction	*CD40*
ENSG00000101017	GO:0006954	inflammatory response	*CD40*
ENSG00000101017	GO:0045944	positive regulation of transcription from RNA polymerase II promoter	*CD40*
ENSG00000101017	GO:0043123	positive regulation of I-kappaB kinase/NF-kappaB cascade	*CD40*
ENSG00000101017	GO:0006874	cellular calcium ion homeostasis	*CD40*
ENSG00000101017	GO:0006461	protein complex assembly	*CD40*
ENSG00000101017	GO:0043406	positive regulation of MAP kinase activity	*CD40*
ENSG00000101017	GO:0030168	platelet activation	*CD40*
ENSG00000101017	GO:0051023	regulation of immunoglobulin secretion	*CD40*
ENSG00000101017	GO:0051092	positive regulation of NF-kappaB transcription factor activity	*CD40*
ENSG00000101017	GO:0001934	positive regulation of protein phosphorylation	*CD40*
ENSG00000101017	GO:0051607	defense response to virus	*CD40*
ENSG00000101017	GO:0030890	positive regulation of B cell proliferation	*CD40*
ENSG00000101017	GO:0042100	B cell proliferation	*CD40*
ENSG00000101017	GO:0071260	cellular response to mechanical stimulus	*CD40*
ENSG00000101017	GO:0032735	positive regulation of interleukin-12 production	*CD40*
ENSG00000101017	GO:0002768	immune response-regulating cell surface receptor signaling pathway	*CD40*
ENSG00000101017	GO:0048304	positive regulation of isotype switching to IgG isotypes	*CD40*
ENSG00000101017	GO:0032855	positive regulation of Rac GTPase activity	*CD40*
ENSG00000101017	GO:0090037	positive regulation of protein kinase C signaling cascade	*CD40*
ENSG00000101017	GO:0042511	positive regulation of tyrosine phosphorylation of Stat1 protein	*CD40*
ENSG00000101017	GO:0043089	positive regulation of Cdc42 GTPase activity	*CD40*
ENSG00000101017	GO:2000353	positive regulation of endothelial cell apoptosis	*CD40*
ENSG00000101017	GO:0042113	B cell activation	*CD40*
ENSG00000101017	GO:0005886	plasma membrane	*CD40*
ENSG00000101017	GO:0005576	extracellular region	*CD40*

ENSG00000101017	GO:0043231	intracellular membrane-bounded organelle	*CD40*
ENSG00000101017	GO:0005887	integral to plasma membrane	*CD40*
ENSG00000101017	GO:0009897	external side of plasma membrane	*CD40*
ENSG00000101017	GO:0035631	CD40 receptor complex	*CD40*
ENSG00000101017	GO:0004872	receptor activity	*CD40*
ENSG00000101017	GO:0005488	binding	*CD40*
ENSG00000101017	GO:0005515	protein binding	*CD40*
ENSG00000101017	GO:0004871	signal transducer activity	*CD40*
ENSG00000101017	GO:0019899	enzyme binding	*CD40*
ENSG00000154162	GO:0007156	homophilic cell adhesion	*CDH12*
ENSG00000154162	GO:0016337	cell-cell adhesion	*CDH12*
ENSG00000154162	GO:0034332	adherens junction organization	*CDH12*
ENSG00000154162	GO:0034329	cell junction assembly	*CDH12*
ENSG00000154162	GO:0045216	cell-cell junction organization	*CDH12*
ENSG00000154162	GO:0016020	membrane	*CDH12*
ENSG00000154162	GO:0016021	integral to membrane	*CDH12*
ENSG00000154162	GO:0005886	plasma membrane	*CDH12*
ENSG00000154162	GO:0005509	calcium ion binding	*CDH12*
ENSG00000154162	GO:0007155	cell adhesion	*CDH12*
ENSG00000150394	GO:0007155	cell adhesion	*CDH8*
ENSG00000150394	GO:0007156	homophilic cell adhesion	*CDH8*
ENSG00000150394	GO:0034332	adherens junction organization	*CDH8*
ENSG00000150394	GO:0034329	cell junction assembly	*CDH8*
ENSG00000150394	GO:0045216	cell-cell junction organization	*CDH8*
ENSG00000150394	GO:0016020	membrane	*CDH8*
ENSG00000150394	GO:0016021	integral to membrane	*CDH8*
ENSG00000150394	GO:0005886	plasma membrane	*CDH8*
ENSG00000150394	GO:0005509	calcium ion binding	*CDH8*
ENSG00000184984	GO:0008283	cell proliferation	*CHRM5*
ENSG00000184984	GO:0007186	G-protein coupled receptor signaling pathway	*CHRM5*
ENSG00000184984	GO:0007213	G-protein coupled acetylcholine receptor signaling pathway	*CHRM5*
ENSG00000184984	GO:0019226	transmission of nerve impulse	*CHRM5*
ENSG00000184984	GO:0007197	inhibition of adenylate cyclase activity by G-protein coupled acetylcholine receptor signaling pathway	*CHRM5*
ENSG00000184984	GO:0015872	dopamine transport	*CHRM5*
ENSG00000184984	GO:0060304	regulation of phosphatidylinositol dephosphorylation	*CHRM5*
ENSG00000184984	GO:0016021	integral to membrane	*CHRM5*
ENSG00000184984	GO:0005886	plasma membrane	*CHRM5*
ENSG00000184984	GO:0030054	cell junction	*CHRM5*
ENSG00000184984	GO:0045202	synapse	*CHRM5*
ENSG00000184984	GO:0045211	postsynaptic membrane	*CHRM5*
ENSG00000184984	GO:0005887	integral to plasma membrane	*CHRM5*
ENSG00000184984	GO:0004872	receptor activity	*CHRM5*
ENSG00000184984	GO:0004871	signal transducer activity	*CHRM5*
ENSG00000184984	GO:0004930	G-protein coupled receptor activity	*CHRM5*
ENSG00000184984	GO:0004435	phosphatidylinositol phospholipase C activity	*CHRM5*
ENSG00000184984	GO:0016907	G-protein coupled acetylcholine receptor activity	*CHRM5*
ENSG00000152910	GO:0007165	signal transduction	*CNTNAP4*
ENSG00000152910	GO:0007155	cell adhesion	*CNTNAP4*
ENSG00000152910	GO:0030425	dendrite	*CNTNAP4*
ENSG00000152910	GO:0005515	protein binding	*CNTNAP4*

ENSG00000152910	GO:0005102	receptor binding	*CNTNAP4*
ENSG00000152910	GO:0016020	membrane	*CNTNAP4*
ENSG00000152910	GO:0016021	integral to membrane	*CNTNAP4*
ENSG00000188517	GO:0016020	membrane	*COL25A1*
ENSG00000188517	GO:0016021	integral to membrane	*COL25A1*
ENSG00000188517	GO:0005615	extracellular space	*COL25A1*
ENSG00000188517	GO:0005581	collagen	*COL25A1*
ENSG00000188517	GO:0001540	beta-amyloid binding	*COL25A1*
ENSG00000188517	GO:0008201	heparin binding	*COL25A1*
ENSG00000144810	GO:0007155	cell adhesion	*COL8A1*
ENSG00000144810	GO:0001525	angiogenesis	*COL8A1*
ENSG00000144810	GO:0050673	epithelial cell proliferation	*COL8A1*
ENSG00000144810	GO:0010811	positive regulation of cell-substrate adhesion	*COL8A1*
ENSG00000144810	GO:0048593	camera-type eye morphogenesis	*COL8A1*
ENSG00000144810	GO:0005576	extracellular region	*COL8A1*
ENSG00000144810	GO:0005581	collagen	*COL8A1*
ENSG00000144810	GO:0005604	basement membrane	*COL8A1*
ENSG00000144810	GO:0031012	extracellular matrix	*COL8A1*
ENSG00000144810	GO:0005591	collagen type VIII	*COL8A1*
ENSG00000100368	GO:0007165	signal transduction	*CSF2RB*
ENSG00000100368	GO:0019221	cytokine-mediated signaling pathway	*CSF2RB*
ENSG00000100368	GO:0007585	respiratory gaseous exchange	*CSF2RB*
ENSG00000100368	GO:0036016	cellular response to interleukin-3	*CSF2RB*
ENSG00000100368	GO:0038043	interleukin-5-mediated signaling pathway	*CSF2RB*
ENSG00000100368	GO:0016021	integral to membrane	*CSF2RB*
ENSG00000100368	GO:0005886	plasma membrane	*CSF2RB*
ENSG00000100368	GO:0005887	integral to plasma membrane	*CSF2RB*
ENSG00000100368	GO:0030526	granulocyte macrophage colony-stimulating factor receptor complex	*CSF2RB*
ENSG00000100368	GO:0004872	receptor activity	*CSF2RB*
ENSG00000100368	GO:0005515	protein binding	*CSF2RB*
ENSG00000100368	GO:0004896	cytokine receptor activity	*CSF2RB*
ENSG00000100368	GO:0004912	interleukin-3 receptor activity	*CSF2RB*
ENSG00000100368	GO:0004914	interleukin-5 receptor activity	*CSF2RB*
ENSG00000100368	GO:0016020	membrane	*CSF2RB*
ENSG00000173406	GO:0021942	radial glia guided migration of Purkinje cell	*DAB1*
ENSG00000173406	GO:0021813	cell-cell adhesion involved in neuronal-glial interactions involved in cerebral cortex radial glia guided migration	*DAB1*
ENSG00000173406	GO:0021799	cerebral cortex radially oriented cell migration	*DAB1*
ENSG00000173406	GO:0021589	cerebellum structural organization	*DAB1*
ENSG00000173406	GO:0021795	cerebral cortex cell migration	*DAB1*
ENSG00000173406	GO:0046426	negative regulation of JAK-STAT cascade	*DAB1*
ENSG00000173406	GO:0048712	negative regulation of astrocyte differentiation	*DAB1*
ENSG00000173406	GO:0007162	negative regulation of cell adhesion	*DAB1*
ENSG00000173406	GO:0016358	dendrite development	*DAB1*
ENSG00000173406	GO:0045666	positive regulation of neuron differentiation	*DAB1*
ENSG00000173406	GO:0021517	ventral spinal cord development	*DAB1*
ENSG00000173406	GO:0050771	negative regulation of axonogenesis	*DAB1*
ENSG00000173406	GO:0051645	Golgi localization	*DAB1*
ENSG00000173406	GO:0045860	positive regulation of protein kinase activity	*DAB1*
ENSG00000173406	GO:0001764	neuron migration	*DAB1*
ENSG00000173406	GO:0007628	adult walking behavior	*DAB1*
ENSG00000173406	GO:0007264	small GTPase mediated signal transduction	*DAB1*

ENSG00000173406	GO:0048471	perinuclear region of cytoplasm	*DAB1*
ENSG00000173406	GO:0005515	protein binding	*DAB1*
ENSG00000173406	GO:0007275	multicellular organismal development	*DAB1*
ENSG00000173406	GO:0030154	cell differentiation	*DAB1*
ENSG00000173406	GO:0007399	nervous system development	*DAB1*
ENSG00000164934	GO:0016567	protein ubiquitination	*DCAF13*
ENSG00000164934	GO:0006364	rRNA processing	*DCAF13*
ENSG00000164934	GO:0005634	nucleus	*DCAF13*
ENSG00000164934	GO:0005730	nucleolus	*DCAF13*
ENSG00000164934	GO:0030529	ribonucleoprotein complex	*DCAF13*
ENSG00000164934	GO:0080008	CUL4 RING ubiquitin ligase complex	*DCAF13*
ENSG00000164934	GO:0005515	protein binding	*DCAF13*
ENSG00000204843	GO:0007399	nervous system development	*DCTN1*
ENSG00000204843	GO:0007067	mitosis	*DCTN1*
ENSG00000204843	GO:0000278	mitotic cell cycle	*DCTN1*
ENSG00000204843	GO:0006810	transport	*DCTN1*
ENSG00000204843	GO:0000086	G2/M transition of mitotic cell cycle	*DCTN1*
ENSG00000204843	GO:0008219	cell death	*DCTN1*
ENSG00000204843	GO:0030968	endoplasmic reticulum unfolded protein response	*DCTN1*
ENSG00000204843	GO:0006987	activation of signaling protein activity involved in unfolded protein response	*DCTN1*
ENSG00000204843	GO:0005737	cytoplasm	*DCTN1*
ENSG00000204843	GO:0005829	cytosol	*DCTN1*
ENSG00000204843	GO:0005856	cytoskeleton	*DCTN1*
ENSG00000204843	GO:0005813	centrosome	*DCTN1*
ENSG00000204843	GO:0005874	microtubule	*DCTN1*
ENSG00000204843	GO:0000776	kinetochore	*DCTN1*
ENSG00000204843	GO:0031252	cell leading edge	*DCTN1*
ENSG00000204843	GO:0000922	spindle pole	*DCTN1*
ENSG00000204843	GO:0005868	cytoplasmic dynein complex	*DCTN1*
ENSG00000204843	GO:0005515	protein binding	*DCTN1*
ENSG00000204843	GO:0003774	motor activity	*DCTN1*
ENSG00000185842	GO:0007018	microtubule-based movement	*DNAH14*
ENSG00000185842	GO:0005737	cytoplasm	*DNAH14*
ENSG00000185842	GO:0005856	cytoskeleton	*DNAH14*
ENSG00000185842	GO:0030286	dynein complex	*DNAH14*
ENSG00000185842	GO:0005874	microtubule	*DNAH14*
ENSG00000185842	GO:0005929	cilium	*DNAH14*
ENSG00000185842	GO:0035085	cilium axoneme	*DNAH14*
ENSG00000185842	GO:0000166	nucleotide binding	*DNAH14*
ENSG00000185842	GO:0005524	ATP binding	*DNAH14*
ENSG00000185842	GO:0003777	microtubule motor activity	*DNAH14*
ENSG00000185842	GO:0017111	nucleoside-triphosphatase activity	*DNAH14*
ENSG00000107099	GO:0007596	blood coagulation	*DOCK8*
ENSG00000107099	GO:0005829	cytosol	*DOCK8*
ENSG00000107099	GO:0005488	binding	*DOCK8*
ENSG00000107099	GO:0005525	GTP binding	*DOCK8*
ENSG00000107099	GO:0005085	guanyl-nucleotide exchange factor activity	*DOCK8*
ENSG00000107099	GO:0051020	GTPase binding	*DOCK8*
ENSG00000171587	GO:0007399	nervous system development	*DSCAM*
ENSG00000171587	GO:0007626	locomotory behavior	*DSCAM*
ENSG00000171587	GO:0007155	cell adhesion	*DSCAM*
ENSG00000171587	GO:0048813	dendrite morphogenesis	*DSCAM*
ENSG00000171587	GO:0007162	negative regulation of cell adhesion	*DSCAM*

ENSG00000171587	GO:0042327	positive regulation of phosphorylation	*DSCAM*
ENSG00000171587	GO:0060060	post-embryonic retina morphogenesis in camera-type eye	*DSCAM*
ENSG00000171587	GO:0048842	positive regulation of axon extension involved in axon guidance	*DSCAM*
ENSG00000171587	GO:0070593	dendrite self-avoidance	*DSCAM*
ENSG00000171587	GO:0005886	plasma membrane	*DSCAM*
ENSG00000171587	GO:0005576	extracellular region	*DSCAM*
ENSG00000171587	GO:0030424	axon	*DSCAM*
ENSG00000171587	GO:0005624	membrane fraction	*DSCAM*
ENSG00000171587	GO:0005887	integral to plasma membrane	*DSCAM*
ENSG00000171587	GO:0030426	growth cone	*DSCAM*
ENSG00000171587	GO:0005515	protein binding	*DSCAM*
ENSG00000203965	GO:0005509	calcium ion binding	*EFCAB7*
ENSG00000080224	GO:0006468	protein phosphorylation	*EPHA6*
ENSG00000080224	GO:0016020	membrane	*EPHA6*
ENSG00000080224	GO:0016021	integral to membrane	*EPHA6*
ENSG00000080224	GO:0005515	protein binding	*EPHA6*
ENSG00000080224	GO:0005524	ATP binding	*EPHA6*
ENSG00000080224	GO:0005003	ephrin receptor activity	*EPHA6*
ENSG00000080224	GO:0008150	biological_process	*EPHA6*
ENSG00000080224	GO:0005575	cellular_component	*EPHA6*
ENSG00000080224	GO:0005887	integral to plasma membrane	*EPHA6*
ENSG00000080224	GO:0004872	receptor activity	*EPHA6*
ENSG00000080224	GO:0000166	nucleotide binding	*EPHA6*
ENSG00000080224	GO:0003674	molecular_function	*EPHA6*
ENSG00000080224	GO:0004672	protein kinase activity	*EPHA6*
ENSG00000080224	GO:0004713	protein tyrosine kinase activity	*EPHA6*
ENSG00000080224	GO:0016772	transferase activity, transferring phosphorus-containing groups	*EPHA6*
ENSG00000178568	GO:0008283	cell proliferation	*ERBB4*
ENSG00000178568	GO:0007275	multicellular organismal development	*ERBB4*
ENSG00000178568	GO:0006915	apoptotic process	*ERBB4*
ENSG00000178568	GO:0007399	nervous system development	*ERBB4*
ENSG00000178568	GO:0008285	negative regulation of cell proliferation	*ERBB4*
ENSG00000178568	GO:0021537	telencephalon development	*ERBB4*
ENSG00000178568	GO:0045893	positive regulation of transcription, DNA-dependent	*ERBB4*
ENSG00000178568	GO:0050679	positive regulation of epithelial cell proliferation	*ERBB4*
ENSG00000178568	GO:0008284	positive regulation of cell proliferation	*ERBB4*
ENSG00000178568	GO:0007165	signal transduction	*ERBB4*
ENSG00000178568	GO:0043066	negative regulation of apoptotic process	*ERBB4*
ENSG00000178568	GO:0045165	cell fate commitment	*ERBB4*
ENSG00000178568	GO:0007507	heart development	*ERBB4*
ENSG00000178568	GO:0018108	peptidyl-tyrosine phosphorylation	*ERBB4*
ENSG00000178568	GO:0060749	mammary gland alveolus development	*ERBB4*
ENSG00000178568	GO:0006468	protein phosphorylation	*ERBB4*
ENSG00000178568	GO:0007169	transmembrane receptor protein tyrosine kinase signaling pathway	*ERBB4*
ENSG00000178568	GO:0060045	positive regulation of cardiac muscle cell proliferation	*ERBB4*
ENSG00000178568	GO:0042493	response to drug	*ERBB4*
ENSG00000178568	GO:0030334	regulation of cell migration	*ERBB4*
ENSG00000178568	GO:0001755	neural crest cell migration	*ERBB4*
ENSG00000178568	GO:0032570	response to progesterone stimulus	*ERBB4*

ENSG00000178568	GO:0070374	positive regulation of ERK1 and ERK2 cascade	*ERBB4*
ENSG00000178568	GO:0009880	embryonic pattern specification	*ERBB4*
ENSG00000178568	GO:0030335	positive regulation of cell migration	*ERBB4*
ENSG00000178568	GO:0046326	positive regulation of glucose import	*ERBB4*
ENSG00000178568	GO:0042542	response to hydrogen peroxide	*ERBB4*
ENSG00000178568	GO:0045768	positive regulation of anti-apoptosis	*ERBB4*
ENSG00000178568	GO:0060644	mammary gland epithelial cell differentiation	*ERBB4*
ENSG00000178568	GO:0046777	protein autophosphorylation	*ERBB4*
ENSG00000178568	GO:0001934	positive regulation of protein phosphorylation	*ERBB4*
ENSG00000178568	GO:0007595	lactation	*ERBB4*
ENSG00000178568	GO:0016477	cell migration	*ERBB4*
ENSG00000178568	GO:0060074	synapse maturation	*ERBB4*
ENSG00000178568	GO:0032230	positive regulation of synaptic transmission, GABAergic	*ERBB4*
ENSG00000178568	GO:0043129	surfactant homeostasis	*ERBB4*
ENSG00000178568	GO:0043552	positive regulation of phosphatidylinositol 3-kinase activity	*ERBB4*
ENSG00000178568	GO:0021551	central nervous system morphogenesis	*ERBB4*
ENSG00000178568	GO:0021889	olfactory bulb interneuron differentiation	*ERBB4*
ENSG00000178568	GO:0042523	positive regulation of tyrosine phosphorylation of Stat5 protein	*ERBB4*
ENSG00000178568	GO:0061026	cardiac muscle tissue regeneration	*ERBB4*
ENSG00000178568	GO:2000366	positive regulation of STAT protein import into nucleus	*ERBB4*
ENSG00000178568	GO:0030879	mammary gland development	*ERBB4*
ENSG00000178568	GO:0016020	membrane	*ERBB4*
ENSG00000178568	GO:0016021	integral to membrane	*ERBB4*
ENSG00000178568	GO:0005634	nucleus	*ERBB4*
ENSG00000178568	GO:0005737	cytoplasm	*ERBB4*
ENSG00000178568	GO:0005886	plasma membrane	*ERBB4*
ENSG00000178568	GO:0005576	extracellular region	*ERBB4*
ENSG00000178568	GO:0005739	mitochondrion	*ERBB4*
ENSG00000178568	GO:0005829	cytosol	*ERBB4*
ENSG00000178568	GO:0005759	mitochondrial matrix	*ERBB4*
ENSG00000178568	GO:0016323	basolateral plasma membrane	*ERBB4*
ENSG00000178568	GO:0005654	nucleoplasm	*ERBB4*
ENSG00000178568	GO:0014069	postsynaptic density	*ERBB4*
ENSG00000178568	GO:0005901	caveola	*ERBB4*
ENSG00000178568	GO:0045121	membrane raft	*ERBB4*
ENSG00000178568	GO:0004872	receptor activity	*ERBB4*
ENSG00000178568	GO:0000166	nucleotide binding	*ERBB4*
ENSG00000178568	GO:0042803	protein homodimerization activity	*ERBB4*
ENSG00000178568	GO:0005515	protein binding	*ERBB4*
ENSG00000178568	GO:0005524	ATP binding	*ERBB4*
ENSG00000178568	GO:0005154	epidermal growth factor receptor binding	*ERBB4*
ENSG00000178568	GO:0004672	protein kinase activity	*ERBB4*
ENSG00000178568	GO:0004714	transmembrane receptor protein tyrosine kinase activity	*ERBB4*
ENSG00000178568	GO:0004713	protein tyrosine kinase activity	*ERBB4*
ENSG00000178568	GO:0044212	transcription regulatory region DNA binding	*ERBB4*
ENSG00000178568	GO:0004716	receptor signaling protein tyrosine kinase activity	*ERBB4*
ENSG00000143971	GO:0005634	nucleus	*ETAA1*
ENSG00000143971	GO:0005737	cytoplasm	*ETAA1*
ENSG00000188107	GO:0005515	protein binding	*EYS*

ENSG00000188107	GO:0005509	calcium ion binding	*EYS*
ENSG00000188107	GO:0007601	visual perception	*EYS*
ENSG00000188107	GO:0050896	response to stimulus	*EYS*
ENSG00000188107	GO:0005576	extracellular region	*EYS*
ENSG00000148541			*FAM13C*
ENSG00000197872	GO:0005622	intracellular	*FAM49A*
ENSG00000170613	GO:0005634	nucleus	*FAM71B*
ENSG00000170613	GO:0005737	cytoplasm	*FAM71B*
ENSG00000026103	GO:0006915	apoptotic process	*FAS*
ENSG00000026103	GO:0007165	signal transduction	*FAS*
ENSG00000026103	GO:0042981	regulation of apoptotic process	*FAS*
ENSG00000026103	GO:0007283	spermatogenesis	*FAS*
ENSG00000026103	GO:0048536	spleen development	*FAS*
ENSG00000026103	GO:0006917	induction of apoptosis	*FAS*
ENSG00000026103	GO:0006955	immune response	*FAS*
ENSG00000026103	GO:0010467	gene expression	*FAS*
ENSG00000026103	GO:0032496	response to lipopolysaccharide	*FAS*
ENSG00000026103	GO:0043627	response to estrogen stimulus	*FAS*
ENSG00000026103	GO:0042493	response to drug	*FAS*
ENSG00000026103	GO:0043065	positive regulation of apoptotic process	*FAS*
ENSG00000026103	GO:0042698	ovulation cycle	*FAS*
ENSG00000026103	GO:0007568	aging	*FAS*
ENSG00000026103	GO:0009636	response to toxin	*FAS*
ENSG00000026103	GO:0034097	response to cytokine stimulus	*FAS*
ENSG00000026103	GO:0006461	protein complex assembly	*FAS*
ENSG00000026103	GO:0001666	response to hypoxia	*FAS*
ENSG00000026103	GO:0042127	regulation of cell proliferation	*FAS*
ENSG00000026103	GO:0008625	induction of apoptosis via death domain receptors	*FAS*
ENSG00000026103	GO:0010940	positive regulation of necrotic cell death	*FAS*
ENSG00000026103	GO:0006916	anti-apoptosis	*FAS*
ENSG00000026103	GO:0051384	response to glucocorticoid stimulus	*FAS*
ENSG00000026103	GO:0010035	response to inorganic substance	*FAS*
ENSG00000026103	GO:0045060	negative thymic T cell selection	*FAS*
ENSG00000026103	GO:0051402	neuron apoptosis	*FAS*
ENSG00000026103	GO:0019724	B cell mediated immunity	*FAS*
ENSG00000026103	GO:0043434	response to peptide hormone stimulus	*FAS*
ENSG00000026103	GO:0006919	activation of cysteine-type endopeptidase activity involved in apoptotic process	*FAS*
ENSG00000026103	GO:0003014	renal system process	*FAS*
ENSG00000026103	GO:0071285	cellular response to lithium ion	*FAS*
ENSG00000026103	GO:0008633	activation of pro-apoptotic gene products	*FAS*
ENSG00000026103	GO:0071260	cellular response to mechanical stimulus	*FAS*
ENSG00000026103	GO:0008624	induction of apoptosis by extracellular signals	*FAS*
ENSG00000026103	GO:0070848	response to growth factor stimulus	*FAS*
ENSG00000026103	GO:0032464	positive regulation of protein homooligomerization	*FAS*
ENSG00000026103	GO:0006924	activation-induced cell death of T cells	*FAS*
ENSG00000026103	GO:0006927	transformed cell apoptosis	*FAS*
ENSG00000026103	GO:0006925	inflammatory cell apoptosis	*FAS*
ENSG00000026103	GO:0045637	regulation of myeloid cell differentiation	*FAS*
ENSG00000026103	GO:0050869	negative regulation of B cell activation	*FAS*
ENSG00000026103	GO:0002377	immunoglobulin production	*FAS*
ENSG00000026103	GO:0031104	dendrite regeneration	*FAS*
ENSG00000026103	GO:0045619	regulation of lymphocyte differentiation	*FAS*

ENSG00000026103	GO:0071455	cellular response to hyperoxia	*FAS*
ENSG00000026103	GO:0033209	tumor necrosis factor-mediated signaling pathway	*FAS*
ENSG00000026103	GO:0043066	negative regulation of apoptotic process	*FAS*
ENSG00000026103	GO:0043029	T cell homeostasis	*FAS*
ENSG00000026103	GO:0051260	protein homooligomerization	*FAS*
ENSG00000026103	GO:0016021	integral to membrane	*FAS*
ENSG00000026103	GO:0005634	nucleus	*FAS*
ENSG00000026103	GO:0005737	cytoplasm	*FAS*
ENSG00000026103	GO:0005886	plasma membrane	*FAS*
ENSG00000026103	GO:0005576	extracellular region	*FAS*
ENSG00000026103	GO:0005829	cytosol	*FAS*
ENSG00000026103	GO:0048471	perinuclear region of cytoplasm	*FAS*
ENSG00000026103	GO:0005625	soluble fraction	*FAS*
ENSG00000026103	GO:0045121	membrane raft	*FAS*
ENSG00000026103	GO:0009897	external side of plasma membrane	*FAS*
ENSG00000026103	GO:0031265	CD95 death-inducing signaling complex	*FAS*
ENSG00000026103	GO:0031264	death-inducing signaling complex	*FAS*
ENSG00000026103	GO:0009986	cell surface	*FAS*
ENSG00000026103	GO:0004872	receptor activity	*FAS*
ENSG00000026103	GO:0005488	binding	*FAS*
ENSG00000026103	GO:0005515	protein binding	*FAS*
ENSG00000026103	GO:0019900	kinase binding	*FAS*
ENSG00000026103	GO:0004871	signal transducer activity	*FAS*
ENSG00000026103	GO:0004888	transmembrane signaling receptor activity	*FAS*
ENSG00000026103	GO:0042802	identical protein binding	*FAS*
ENSG00000026103	GO:0005031	tumor necrosis factor-activated receptor activity	*FAS*
ENSG00000176988	GO:0016020	membrane	*FMR1NB*
ENSG00000176988	GO:0016021	integral to membrane	*FMR1NB*
ENSG00000164694	GO:0071456	cellular response to hypoxia	*FNDC1*
ENSG00000164694	GO:0001934	positive regulation of protein phosphorylation	*FNDC1*
ENSG00000164694	GO:0051223	regulation of protein transport	*FNDC1*
ENSG00000164694	GO:0010666	positive regulation of cardiac muscle cell apoptosis	*FNDC1*
ENSG00000164694	GO:0005737	cytoplasm	*FNDC1*
ENSG00000164694	GO:0005886	plasma membrane	*FNDC1*
ENSG00000164694	GO:0005576	extracellular region	*FNDC1*
ENSG00000164694	GO:0031966	mitochondrial membrane	*FNDC1*
ENSG00000164694	GO:0005911	cell-cell junction	*FNDC1*
ENSG00000164694	GO:0005515	protein binding	*FNDC1*
ENSG00000164930	GO:0007275	multicellular organismal development	*FZD6*
ENSG00000164930	GO:0016055	Wnt receptor signaling pathway	*FZD6*
ENSG00000164930	GO:0001736	establishment of planar polarity	*FZD6*
ENSG00000164930	GO:0007409	axonogenesis	*FZD6*
ENSG00000164930	GO:0042472	inner ear morphogenesis	*FZD6*
ENSG00000164930	GO:0001525	angiogenesis	*FZD6*
ENSG00000164930	GO:0043433	negative regulation of sequence-specific DNA binding transcription factor activity	*FZD6*
ENSG00000164930	GO:0001843	neural tube closure	*FZD6*
ENSG00000164930	GO:0090090	negative regulation of canonical Wnt receptor signaling pathway	*FZD6*
ENSG00000164930	GO:0007166	cell surface receptor signaling pathway	*FZD6*
ENSG00000164930	GO:0030168	platelet activation	*FZD6*
ENSG00000164930	GO:0001942	hair follicle development	*FZD6*
ENSG00000164930	GO:0035567	non-canonical Wnt receptor signaling	*FZD6*

		pathway	
ENSG00000164930	GO:0033278	cell proliferation in midbrain	*FZD6*
ENSG00000164930	GO:0060071	Wnt receptor signaling pathway, planar cell polarity pathway	*FZD6*
ENSG00000164930	GO:0008406	gonad development	*FZD6*
ENSG00000164930	GO:0007199	G-protein signaling, coupled to cGMP nucleotide second messenger	*FZD6*
ENSG00000164930	GO:0030901	midbrain development	*FZD6*
ENSG00000164930	GO:0016020	membrane	*FZD6*
ENSG00000164930	GO:0016021	integral to membrane	*FZD6*
ENSG00000164930	GO:0005737	cytoplasm	*FZD6*
ENSG00000164930	GO:0005886	plasma membrane	*FZD6*
ENSG00000164930	GO:0005887	integral to plasma membrane	*FZD6*
ENSG00000164930	GO:0045177	apical part of cell	*FZD6*
ENSG00000164930	GO:0016327	apicolateral plasma membrane	*FZD6*
ENSG00000164930	GO:0032589	neuron projection membrane	*FZD6*
ENSG00000164930	GO:0004872	receptor activity	*FZD6*
ENSG00000164930	GO:0005515	protein binding	*FZD6*
ENSG00000164930	GO:0004871	signal transducer activity	*FZD6*
ENSG00000164930	GO:0004930	G-protein coupled receptor activity	*FZD6*
ENSG00000164930	GO:0004888	transmembrane signaling receptor activity	*FZD6*
ENSG00000164930	GO:0017147	Wnt-protein binding	*FZD6*
ENSG00000164930	GO:0042813	Wnt-activated receptor activity	*FZD6*
ENSG00000164930	GO:0030165	PDZ domain binding	*FZD6*
ENSG00000237051	GO:0008285	negative regulation of cell proliferation	*GABBR1*
ENSG00000237051	GO:0007186	G-protein coupled receptor signaling pathway	*GABBR1*
ENSG00000237051	GO:0001649	osteoblast differentiation	*GABBR1*
ENSG00000237051	GO:0007194	negative regulation of adenylate cyclase activity	*GABBR1*
ENSG00000237051	GO:0014048	regulation of glutamate secretion	*GABBR1*
ENSG00000237051	GO:0014053	negative regulation of gamma-aminobutyric acid secretion	*GABBR1*
ENSG00000237051	GO:0030817	regulation of cAMP biosynthetic process	*GABBR1*
ENSG00000237051	GO:0032811	negative regulation of epinephrine secretion	*GABBR1*
ENSG00000237051	GO:0033602	negative regulation of dopamine secretion	*GABBR1*
ENSG00000237051	GO:0035094	response to nicotine	*GABBR1*
ENSG00000237051	GO:0045471	response to ethanol	*GABBR1*
ENSG00000237051	GO:0050805	negative regulation of synaptic transmission	*GABBR1*
ENSG00000237051	GO:0060124	positive regulation of growth hormone secretion	*GABBR1*
ENSG00000237051	GO:0016020	membrane	*GABBR1*
ENSG00000237051	GO:0016021	integral to membrane	*GABBR1*
ENSG00000237051	GO:0005737	cytoplasm	*GABBR1*
ENSG00000237051	GO:0005789	endoplasmic reticulum membrane	*GABBR1*
ENSG00000237051	GO:0043025	neuronal cell body	*GABBR1*
ENSG00000237051	GO:0019717	synaptosome	*GABBR1*
ENSG00000237051	GO:0008021	synaptic vesicle	*GABBR1*
ENSG00000237051	GO:0031966	mitochondrial membrane	*GABBR1*
ENSG00000237051	GO:0043197	dendritic spine	*GABBR1*
ENSG00000237051	GO:0005625	soluble fraction	*GABBR1*
ENSG00000237051	GO:0005792	microsome	*GABBR1*
ENSG00000237051	GO:0030673	axolemma	*GABBR1*
ENSG00000237051	GO:0043198	dendritic shaft	*GABBR1*
ENSG00000237051	GO:0045121	membrane raft	*GABBR1*
ENSG00000237051	GO:0004872	receptor activity	*GABBR1*

ENSG00000237051	GO:0004871	signal transducer activity	*GABBR1*
ENSG00000237051	GO:0004930	G-protein coupled receptor activity	*GABBR1*
ENSG00000237051	GO:0004965	G-protein coupled GABA receptor activity	*GABBR1*
ENSG00000237051	GO:0008134	transcription factor binding	*GABBR1*
ENSG00000204681	GO:0008285	negative regulation of cell proliferation	*GABBR1*
ENSG00000204681	GO:0007186	G-protein coupled receptor signaling pathway	*GABBR1*
ENSG00000204681	GO:0001649	osteoblast differentiation	*GABBR1*
ENSG00000204681	GO:0007194	negative regulation of adenylate cyclase activity	*GABBR1*
ENSG00000204681	GO:0014048	regulation of glutamate secretion	*GABBR1*
ENSG00000204681	GO:0014053	negative regulation of gamma-aminobutyric acid secretion	*GABBR1*
ENSG00000204681	GO:0030817	regulation of cAMP biosynthetic process	*GABBR1*
ENSG00000204681	GO:0032811	negative regulation of epinephrine secretion	*GABBR1*
ENSG00000204681	GO:0033602	negative regulation of dopamine secretion	*GABBR1*
ENSG00000204681	GO:0035094	response to nicotine	*GABBR1*
ENSG00000204681	GO:0045471	response to ethanol	*GABBR1*
ENSG00000204681	GO:0050805	negative regulation of synaptic transmission	*GABBR1*
ENSG00000204681	GO:0060124	positive regulation of growth hormone secretion	*GABBR1*
ENSG00000204681	GO:0016020	membrane	*GABBR1*
ENSG00000204681	GO:0016021	integral to membrane	*GABBR1*
ENSG00000204681	GO:0005737	cytoplasm	*GABBR1*
ENSG00000204681	GO:0005789	endoplasmic reticulum membrane	*GABBR1*
ENSG00000204681	GO:0043025	neuronal cell body	*GABBR1*
ENSG00000204681	GO:0019717	synaptosome	*GABBR1*
ENSG00000204681	GO:0008021	synaptic vesicle	*GABBR1*
ENSG00000204681	GO:0031966	mitochondrial membrane	*GABBR1*
ENSG00000204681	GO:0043197	dendritic spine	*GABBR1*
ENSG00000204681	GO:0005625	soluble fraction	*GABBR1*
ENSG00000204681	GO:0005792	microsome	*GABBR1*
ENSG00000204681	GO:0030673	axolemma	*GABBR1*
ENSG00000204681	GO:0043198	dendritic shaft	*GABBR1*
ENSG00000204681	GO:0045121	membrane raft	*GABBR1*
ENSG00000204681	GO:0004872	receptor activity	*GABBR1*
ENSG00000204681	GO:0004871	signal transducer activity	*GABBR1*
ENSG00000204681	GO:0004930	G-protein coupled receptor activity	*GABBR1*
ENSG00000204681	GO:0004965	G-protein coupled GABA receptor activity	*GABBR1*
ENSG00000204681	GO:0008134	transcription factor binding	*GABBR1*
ENSG00000204681	GO:0007268	synaptic transmission	*GABBR1*
ENSG00000204681	GO:0007214	gamma-aminobutyric acid signaling pathway	*GABBR1*
ENSG00000204681	GO:0005886	plasma membrane	*GABBR1*
ENSG00000204681	GO:0005576	extracellular region	*GABBR1*
ENSG00000204681	GO:0030054	cell junction	*GABBR1*
ENSG00000204681	GO:0045202	synapse	*GABBR1*
ENSG00000204681	GO:0045211	postsynaptic membrane	*GABBR1*
ENSG00000204681	GO:0005887	integral to plasma membrane	*GABBR1*
ENSG00000204681	GO:0005515	protein binding	*GABBR1*
ENSG00000232569	GO:0008285	negative regulation of cell proliferation	*GABBR1*
ENSG00000232569	GO:0007186	G-protein coupled receptor signaling pathway	*GABBR1*
ENSG00000232569	GO:0001649	osteoblast differentiation	*GABBR1*
ENSG00000232569	GO:0007194	negative regulation of adenylate cyclase activity	*GABBR1*
ENSG00000232569	GO:0014048	regulation of glutamate secretion	*GABBR1*

ENSG00000232569	GO:0014053	negative regulation of gamma-aminobutyric acid secretion	*GABBR1*
ENSG00000232569	GO:0030817	regulation of cAMP biosynthetic process	*GABBR1*
ENSG00000232569	GO:0032811	negative regulation of epinephrine secretion	*GABBR1*
ENSG00000232569	GO:0033602	negative regulation of dopamine secretion	*GABBR1*
ENSG00000232569	GO:0035094	response to nicotine	*GABBR1*
ENSG00000232569	GO:0045471	response to ethanol	*GABBR1*
ENSG00000232569	GO:0050805	negative regulation of synaptic transmission	*GABBR1*
ENSG00000232569	GO:0060124	positive regulation of growth hormone secretion	*GABBR1*
ENSG00000232569	GO:0016020	membrane	*GABBR1*
ENSG00000232569	GO:0016021	integral to membrane	*GABBR1*
ENSG00000232569	GO:0005737	cytoplasm	*GABBR1*
ENSG00000232569	GO:0005789	endoplasmic reticulum membrane	*GABBR1*
ENSG00000232569	GO:0043025	neuronal cell body	*GABBR1*
ENSG00000232569	GO:0019717	synaptosome	*GABBR1*
ENSG00000232569	GO:0008021	synaptic vesicle	*GABBR1*
ENSG00000232569	GO:0031966	mitochondrial membrane	*GABBR1*
ENSG00000232569	GO:0043197	dendritic spine	*GABBR1*
ENSG00000232569	GO:0005625	soluble fraction	*GABBR1*
ENSG00000232569	GO:0005792	microsome	*GABBR1*
ENSG00000232569	GO:0030673	axolemma	*GABBR1*
ENSG00000232569	GO:0043198	dendritic shaft	*GABBR1*
ENSG00000232569	GO:0045121	membrane raft	*GABBR1*
ENSG00000232569	GO:0004872	receptor activity	*GABBR1*
ENSG00000232569	GO:0004871	signal transducer activity	*GABBR1*
ENSG00000232569	GO:0004930	G-protein coupled receptor activity	*GABBR1*
ENSG00000232569	GO:0004965	G-protein coupled GABA receptor activity	*GABBR1*
ENSG00000232569	GO:0008134	transcription factor binding	*GABBR1*
ENSG00000232569	GO:0007268	synaptic transmission	*GABBR1*
ENSG00000232569	GO:0007214	gamma-aminobutyric acid signaling pathway	*GABBR1*
ENSG00000232569	GO:0005886	plasma membrane	*GABBR1*
ENSG00000232569	GO:0005576	extracellular region	*GABBR1*
ENSG00000232569	GO:0030054	cell junction	*GABBR1*
ENSG00000232569	GO:0045202	synapse	*GABBR1*
ENSG00000232569	GO:0045211	postsynaptic membrane	*GABBR1*
ENSG00000232569	GO:0005887	integral to plasma membrane	*GABBR1*
ENSG00000232569	GO:0005515	protein binding	*GABBR1*
ENSG00000206511	GO:0008285	negative regulation of cell proliferation	*GABBR1*
ENSG00000206511	GO:0007186	G-protein coupled receptor signaling pathway	*GABBR1*
ENSG00000206511	GO:0001649	osteoblast differentiation	*GABBR1*
ENSG00000206511	GO:0007194	negative regulation of adenylate cyclase activity	*GABBR1*
ENSG00000206511	GO:0014048	regulation of glutamate secretion	*GABBR1*
ENSG00000206511	GO:0014053	negative regulation of gamma-aminobutyric acid secretion	*GABBR1*
ENSG00000206511	GO:0030817	regulation of cAMP biosynthetic process	*GABBR1*
ENSG00000206511	GO:0032811	negative regulation of epinephrine secretion	*GABBR1*
ENSG00000206511	GO:0033602	negative regulation of dopamine secretion	*GABBR1*
ENSG00000206511	GO:0035094	response to nicotine	*GABBR1*
ENSG00000206511	GO:0045471	response to ethanol	*GABBR1*
ENSG00000206511	GO:0050805	negative regulation of synaptic transmission	*GABBR1*
ENSG00000206511	GO:0060124	positive regulation of growth hormone secretion	*GABBR1*
ENSG00000206511	GO:0016020	membrane	*GABBR1*

ENSG00000206511	GO:0016021	integral to membrane	*GABBR1*
ENSG00000206511	GO:0005737	cytoplasm	*GABBR1*
ENSG00000206511	GO:0005789	endoplasmic reticulum membrane	*GABBR1*
ENSG00000206511	GO:0043025	neuronal cell body	*GABBR1*
ENSG00000206511	GO:0019717	synaptosome	*GABBR1*
ENSG00000206511	GO:0008021	synaptic vesicle	*GABBR1*
ENSG00000206511	GO:0031966	mitochondrial membrane	*GABBR1*
ENSG00000206511	GO:0043197	dendritic spine	*GABBR1*
ENSG00000206511	GO:0005625	soluble fraction	*GABBR1*
ENSG00000206511	GO:0005792	microsome	*GABBR1*
ENSG00000206511	GO:0030673	axolemma	*GABBR1*
ENSG00000206511	GO:0043198	dendritic shaft	*GABBR1*
ENSG00000206511	GO:0045121	membrane raft	*GABBR1*
ENSG00000206511	GO:0004872	receptor activity	*GABBR1*
ENSG00000206511	GO:0004871	signal transducer activity	*GABBR1*
ENSG00000206511	GO:0004930	G-protein coupled receptor activity	*GABBR1*
ENSG00000206511	GO:0004965	G-protein coupled GABA receptor activity	*GABBR1*
ENSG00000206511	GO:0008134	transcription factor binding	*GABBR1*
ENSG00000206511	GO:0007268	synaptic transmission	*GABBR1*
ENSG00000206511	GO:0007214	gamma-aminobutyric acid signaling pathway	*GABBR1*
ENSG00000206511	GO:0005886	plasma membrane	*GABBR1*
ENSG00000206511	GO:0005576	extracellular region	*GABBR1*
ENSG00000206511	GO:0030054	cell junction	*GABBR1*
ENSG00000206511	GO:0045202	synapse	*GABBR1*
ENSG00000206511	GO:0045211	postsynaptic membrane	*GABBR1*
ENSG00000206511	GO:0005887	integral to plasma membrane	*GABBR1*
ENSG00000206511	GO:0005515	protein binding	*GABBR1*
ENSG00000232632	GO:0008285	negative regulation of cell proliferation	*GABBR1*
ENSG00000232632	GO:0007186	G-protein coupled receptor signaling pathway	*GABBR1*
ENSG00000232632	GO:0001649	osteoblast differentiation	*GABBR1*
ENSG00000232632	GO:0007194	negative regulation of adenylate cyclase activity	*GABBR1*
ENSG00000232632	GO:0014048	regulation of glutamate secretion	*GABBR1*
ENSG00000232632	GO:0014053	negative regulation of gamma-aminobutyric acid secretion	*GABBR1*
ENSG00000232632	GO:0030817	regulation of cAMP biosynthetic process	*GABBR1*
ENSG00000232632	GO:0032811	negative regulation of epinephrine secretion	*GABBR1*
ENSG00000232632	GO:0033602	negative regulation of dopamine secretion	*GABBR1*
ENSG00000232632	GO:0035094	response to nicotine	*GABBR1*
ENSG00000232632	GO:0045471	response to ethanol	*GABBR1*
ENSG00000232632	GO:0050805	negative regulation of synaptic transmission	*GABBR1*
ENSG00000232632	GO:0060124	positive regulation of growth hormone secretion	*GABBR1*
ENSG00000232632	GO:0016020	membrane	*GABBR1*
ENSG00000232632	GO:0016021	integral to membrane	*GABBR1*
ENSG00000232632	GO:0005737	cytoplasm	*GABBR1*
ENSG00000232632	GO:0005789	endoplasmic reticulum membrane	*GABBR1*
ENSG00000232632	GO:0043025	neuronal cell body	*GABBR1*
ENSG00000232632	GO:0019717	synaptosome	*GABBR1*
ENSG00000232632	GO:0008021	synaptic vesicle	*GABBR1*
ENSG00000232632	GO:0031966	mitochondrial membrane	*GABBR1*
ENSG00000232632	GO:0043197	dendritic spine	*GABBR1*
ENSG00000232632	GO:0005625	soluble fraction	*GABBR1*
ENSG00000232632	GO:0005792	microsome	*GABBR1*
ENSG00000232632	GO:0030673	axolemma	*GABBR1*

ENSG00000232632	GO:0043198	dendritic shaft	*GABBR1*
ENSG00000232632	GO:0045121	membrane raft	*GABBR1*
ENSG00000232632	GO:0004872	receptor activity	*GABBR1*
ENSG00000232632	GO:0004871	signal transducer activity	*GABBR1*
ENSG00000232632	GO:0004930	G-protein coupled receptor activity	*GABBR1*
ENSG00000232632	GO:0004965	G-protein coupled GABA receptor activity	*GABBR1*
ENSG00000232632	GO:0008134	transcription factor binding	*GABBR1*
ENSG00000232632	GO:0007268	synaptic transmission	*GABBR1*
ENSG00000232632	GO:0007214	gamma-aminobutyric acid signaling pathway	*GABBR1*
ENSG00000232632	GO:0005886	plasma membrane	*GABBR1*
ENSG00000232632	GO:0005576	extracellular region	*GABBR1*
ENSG00000232632	GO:0030054	cell junction	*GABBR1*
ENSG00000232632	GO:0045202	synapse	*GABBR1*
ENSG00000232632	GO:0045211	postsynaptic membrane	*GABBR1*
ENSG00000232632	GO:0005887	integral to plasma membrane	*GABBR1*
ENSG00000232632	GO:0005515	protein binding	*GABBR1*
ENSG00000206466	GO:0008285	negative regulation of cell proliferation	*GABBR1*
ENSG00000206466	GO:0007186	G-protein coupled receptor signaling pathway	*GABBR1*
ENSG00000206466	GO:0001649	osteoblast differentiation	*GABBR1*
ENSG00000206466	GO:0007194	negative regulation of adenylate cyclase activity	*GABBR1*
ENSG00000206466	GO:0014048	regulation of glutamate secretion	*GABBR1*
ENSG00000206466	GO:0014053	negative regulation of gamma-aminobutyric acid secretion	*GABBR1*
ENSG00000206466	GO:0030817	regulation of cAMP biosynthetic process	*GABBR1*
ENSG00000206466	GO:0032811	negative regulation of epinephrine secretion	*GABBR1*
ENSG00000206466	GO:0033602	negative regulation of dopamine secretion	*GABBR1*
ENSG00000206466	GO:0035094	response to nicotine	*GABBR1*
ENSG00000206466	GO:0045471	response to ethanol	*GABBR1*
ENSG00000206466	GO:0050805	negative regulation of synaptic transmission	*GABBR1*
ENSG00000206466	GO:0060124	positive regulation of growth hormone secretion	*GABBR1*
ENSG00000206466	GO:0016020	membrane	*GABBR1*
ENSG00000206466	GO:0016021	integral to membrane	*GABBR1*
ENSG00000206466	GO:0005737	cytoplasm	*GABBR1*
ENSG00000206466	GO:0005789	endoplasmic reticulum membrane	*GABBR1*
ENSG00000206466	GO:0043025	neuronal cell body	*GABBR1*
ENSG00000206466	GO:0019717	synaptosome	*GABBR1*
ENSG00000206466	GO:0008021	synaptic vesicle	*GABBR1*
ENSG00000206466	GO:0031966	mitochondrial membrane	*GABBR1*
ENSG00000206466	GO:0043197	dendritic spine	*GABBR1*
ENSG00000206466	GO:0005625	soluble fraction	*GABBR1*
ENSG00000206466	GO:0005792	microsome	*GABBR1*
ENSG00000206466	GO:0030673	axolemma	*GABBR1*
ENSG00000206466	GO:0043198	dendritic shaft	*GABBR1*
ENSG00000206466	GO:0045121	membrane raft	*GABBR1*
ENSG00000206466	GO:0004872	receptor activity	*GABBR1*
ENSG00000206466	GO:0004871	signal transducer activity	*GABBR1*
ENSG00000206466	GO:0004930	G-protein coupled receptor activity	*GABBR1*
ENSG00000206466	GO:0004965	G-protein coupled GABA receptor activity	*GABBR1*
ENSG00000206466	GO:0008134	transcription factor binding	*GABBR1*
ENSG00000206466	GO:0007268	synaptic transmission	*GABBR1*
ENSG00000206466	GO:0007214	gamma-aminobutyric acid signaling pathway	*GABBR1*
ENSG00000206466	GO:0005886	plasma membrane	*GABBR1*
ENSG00000206466	GO:0005576	extracellular region	*GABBR1*

ENSG00000206466	GO:0030054	cell junction	*GABBR1*
ENSG00000206466	GO:0045202	synapse	*GABBR1*
ENSG00000206466	GO:0045211	postsynaptic membrane	*GABBR1*
ENSG00000206466	GO:0005887	integral to plasma membrane	*GABBR1*
ENSG00000206466	GO:0005515	protein binding	*GABBR1*
ENSG00000237112	GO:0008285	negative regulation of cell proliferation	*GABBR1*
ENSG00000237112	GO:0007186	G-protein coupled receptor signaling pathway	*GABBR1*
ENSG00000237112	GO:0001649	osteoblast differentiation	*GABBR1*
ENSG00000237112	GO:0007194	negative regulation of adenylate cyclase activity	*GABBR1*
ENSG00000237112	GO:0014048	regulation of glutamate secretion	*GABBR1*
ENSG00000237112	GO:0014053	negative regulation of gamma-aminobutyric acid secretion	*GABBR1*
ENSG00000237112	GO:0030817	regulation of cAMP biosynthetic process	*GABBR1*
ENSG00000237112	GO:0032811	negative regulation of epinephrine secretion	*GABBR1*
ENSG00000237112	GO:0033602	negative regulation of dopamine secretion	*GABBR1*
ENSG00000237112	GO:0035094	response to nicotine	*GABBR1*
ENSG00000237112	GO:0045471	response to ethanol	*GABBR1*
ENSG00000237112	GO:0050805	negative regulation of synaptic transmission	*GABBR1*
ENSG00000237112	GO:0060124	positive regulation of growth hormone secretion	*GABBR1*
ENSG00000237112	GO:0016020	membrane	*GABBR1*
ENSG00000237112	GO:0016021	integral to membrane	*GABBR1*
ENSG00000237112	GO:0005737	cytoplasm	*GABBR1*
ENSG00000237112	GO:0005789	endoplasmic reticulum membrane	*GABBR1*
ENSG00000237112	GO:0043025	neuronal cell body	*GABBR1*
ENSG00000237112	GO:0019717	synaptosome	*GABBR1*
ENSG00000237112	GO:0008021	synaptic vesicle	*GABBR1*
ENSG00000237112	GO:0031966	mitochondrial membrane	*GABBR1*
ENSG00000237112	GO:0043197	dendritic spine	*GABBR1*
ENSG00000237112	GO:0005625	soluble fraction	*GABBR1*
ENSG00000237112	GO:0005792	microsome	*GABBR1*
ENSG00000237112	GO:0030673	axolemma	*GABBR1*
ENSG00000237112	GO:0043198	dendritic shaft	*GABBR1*
ENSG00000237112	GO:0045121	membrane raft	*GABBR1*
ENSG00000237112	GO:0004872	receptor activity	*GABBR1*
ENSG00000237112	GO:0004871	signal transducer activity	*GABBR1*
ENSG00000237112	GO:0004930	G-protein coupled receptor activity	*GABBR1*
ENSG00000237112	GO:0004965	G-protein coupled GABA receptor activity	*GABBR1*
ENSG00000237112	GO:0008134	transcription factor binding	*GABBR1*
ENSG00000237112	GO:0007268	synaptic transmission	*GABBR1*
ENSG00000237112	GO:0007214	gamma-aminobutyric acid signaling pathway	*GABBR1*
ENSG00000237112	GO:0005886	plasma membrane	*GABBR1*
ENSG00000237112	GO:0005576	extracellular region	*GABBR1*
ENSG00000237112	GO:0030054	cell junction	*GABBR1*
ENSG00000237112	GO:0045202	synapse	*GABBR1*
ENSG00000237112	GO:0045211	postsynaptic membrane	*GABBR1*
ENSG00000237112	GO:0005887	integral to plasma membrane	*GABBR1*
ENSG00000237112	GO:0005515	protein binding	*GABBR1*
ENSG00000116717	GO:0006915	apoptotic process	*GADD45A*
ENSG00000116717	GO:0006281	DNA repair	*GADD45A*
ENSG00000116717	GO:0006950	response to stress	*GADD45A*
ENSG00000116717	GO:0007050	cell cycle arrest	*GADD45A*
ENSG00000116717	GO:0000086	G2/M transition of mitotic cell cycle	*GADD45A*
ENSG00000116717	GO:0000079	regulation of cyclin-dependent protein kinase	*GADD45A*

		activity	
ENSG00000116717	GO:0006469	negative regulation of protein kinase activity	*GADD45A*
ENSG00000116717	GO:2000379	positive regulation of reactive oxygen species metabolic process	*GADD45A*
ENSG00000116717	GO:0042770	signal transduction in response to DNA damage	*GADD45A*
ENSG00000116717	GO:0007098	centrosome cycle	*GADD45A*
ENSG00000116717	GO:0071260	cellular response to mechanical stimulus	*GADD45A*
ENSG00000116717	GO:0071479	cellular response to ionizing radiation	*GADD45A*
ENSG00000116717	GO:0051726	regulation of cell cycle	*GADD45A*
ENSG00000116717	GO:0005634	nucleus	*GADD45A*
ENSG00000116717	GO:0005515	protein binding	*GADD45A*
ENSG00000141448	GO:0006357	regulation of transcription from RNA polymerase II promoter	*GATA6*
ENSG00000141448	GO:0000122	negative regulation of transcription from RNA polymerase II promoter	*GATA6*
ENSG00000141448	GO:0045893	positive regulation of transcription, DNA-dependent	*GATA6*
ENSG00000141448	GO:0043066	negative regulation of apoptotic process	*GATA6*
ENSG00000141448	GO:0042475	odontogenesis of dentin-containing tooth	*GATA6*
ENSG00000141448	GO:0045944	positive regulation of transcription from RNA polymerase II promoter	*GATA6*
ENSG00000141448	GO:0048645	organ formation	*GATA6*
ENSG00000141448	GO:0045892	negative regulation of transcription, DNA-dependent	*GATA6*
ENSG00000141448	GO:0006355	regulation of transcription, DNA-dependent	*GATA6*
ENSG00000141448	GO:0055007	cardiac muscle cell differentiation	*GATA6*
ENSG00000141448	GO:0060045	positive regulation of cardiac muscle cell proliferation	*GATA6*
ENSG00000141448	GO:0008584	male gonad development	*GATA6*
ENSG00000141448	GO:0043627	response to estrogen stimulus	*GATA6*
ENSG00000141448	GO:0042493	response to drug	*GATA6*
ENSG00000141448	GO:0001701	in utero embryonic development	*GATA6*
ENSG00000141448	GO:0043065	positive regulation of apoptotic process	*GATA6*
ENSG00000141448	GO:0045766	positive regulation of angiogenesis	*GATA6*
ENSG00000141448	GO:0051145	smooth muscle cell differentiation	*GATA6*
ENSG00000141448	GO:0009636	response to toxin	*GATA6*
ENSG00000141448	GO:0051591	response to cAMP	*GATA6*
ENSG00000141448	GO:0071456	cellular response to hypoxia	*GATA6*
ENSG00000141448	GO:0060510	Type II pneumocyte differentiation	*GATA6*
ENSG00000141448	GO:0006644	phospholipid metabolic process	*GATA6*
ENSG00000141448	GO:0001889	liver development	*GATA6*
ENSG00000141448	GO:0060575	intestinal epithelial cell differentiation	*GATA6*
ENSG00000141448	GO:0030513	positive regulation of BMP signaling pathway	*GATA6*
ENSG00000141448	GO:0035239	tube morphogenesis	*GATA6*
ENSG00000141448	GO:0007596	blood coagulation	*GATA6*
ENSG00000141448	GO:0030855	epithelial cell differentiation	*GATA6*
ENSG00000141448	GO:0031016	pancreas development	*GATA6*
ENSG00000141448	GO:0006366	transcription from RNA polymerase II promoter	*GATA6*
ENSG00000141448	GO:0007492	endoderm development	*GATA6*
ENSG00000141448	GO:0003148	outflow tract septum morphogenesis	*GATA6*
ENSG00000141448	GO:0070848	response to growth factor stimulus	*GATA6*
ENSG00000141448	GO:0060430	lung saccule development	*GATA6*
ENSG00000141448	GO:0071158	positive regulation of cell cycle arrest	*GATA6*
ENSG00000141448	GO:0071773	cellular response to BMP stimulus	*GATA6*
ENSG00000141448	GO:0060947	cardiac vascular smooth muscle cell	*GATA6*

		differentiation	
ENSG00000141448	GO:0060486	Clara cell differentiation	*GATA6*
ENSG00000141448	GO:0071371	cellular response to gonadotropin stimulus	*GATA6*
ENSG00000141448	GO:0032911	negative regulation of transforming growth factor beta1 production	*GATA6*
ENSG00000141448	GO:0051891	positive regulation of cardioblast differentiation	*GATA6*
ENSG00000141448	GO:0003309	type B pancreatic cell differentiation	*GATA6*
ENSG00000141448	GO:0003310	pancreatic A cell differentiation	*GATA6*
ENSG00000141448	GO:0007493	endodermal cell fate determination	*GATA6*
ENSG00000141448	GO:0032912	negative regulation of transforming growth factor beta2 production	*GATA6*
ENSG00000141448	GO:0010468	regulation of gene expression	*GATA6*
ENSG00000141448	GO:0006915	apoptotic process	*GATA6*
ENSG00000141448	GO:0005634	nucleus	*GATA6*
ENSG00000141448	GO:0005667	transcription factor complex	*GATA6*
ENSG00000141448	GO:0005654	nucleoplasm	*GATA6*
ENSG00000141448	GO:0003677	DNA binding	*GATA6*
ENSG00000141448	GO:0003700	sequence-specific DNA binding transcription factor activity	*GATA6*
ENSG00000141448	GO:0008270	zinc ion binding	*GATA6*
ENSG00000141448	GO:0005515	protein binding	*GATA6*
ENSG00000141448	GO:0046872	metal ion binding	*GATA6*
ENSG00000141448	GO:0043565	sequence-specific DNA binding	*GATA6*
ENSG00000141448	GO:0000981	sequence-specific DNA binding RNA polymerase II transcription factor activity	*GATA6*
ENSG00000141448	GO:0003690	double-stranded DNA binding	*GATA6*
ENSG00000141448	GO:0003682	chromatin binding	*GATA6*
ENSG00000141448	GO:0019901	protein kinase binding	*GATA6*
ENSG00000141448	GO:0008134	transcription factor binding	*GATA6*
ENSG00000141448	GO:0044212	transcription regulatory region DNA binding	*GATA6*
ENSG00000141448	GO:0003705	sequence-specific distal enhancer binding RNA polymerase II transcription factor activity	*GATA6*
ENSG00000141448	GO:0001103	RNA polymerase II repressing transcription factor binding	*GATA6*
ENSG00000092208	GO:0006397	mRNA processing	*GEMIN2*
ENSG00000092208	GO:0010467	gene expression	*GEMIN2*
ENSG00000092208	GO:0016070	RNA metabolic process	*GEMIN2*
ENSG00000092208	GO:0008380	RNA splicing	*GEMIN2*
ENSG00000092208	GO:0000245	spliceosome assembly	*GEMIN2*
ENSG00000092208	GO:0000375	RNA splicing, via transesterification reactions	*GEMIN2*
ENSG00000092208	GO:0000387	spliceosomal snRNP assembly	*GEMIN2*
ENSG00000092208	GO:0034660	ncRNA metabolic process	*GEMIN2*
ENSG00000092208	GO:0005634	nucleus	*GEMIN2*
ENSG00000092208	GO:0005737	cytoplasm	*GEMIN2*
ENSG00000092208	GO:0005622	intracellular	*GEMIN2*
ENSG00000092208	GO:0005829	cytosol	*GEMIN2*
ENSG00000092208	GO:0005654	nucleoplasm	*GEMIN2*
ENSG00000092208	GO:0005681	spliceosomal complex	*GEMIN2*
ENSG00000092208	GO:0015030	Cajal body	*GEMIN2*
ENSG00000092208	GO:0005515	protein binding	*GEMIN2*
ENSG00000106571	GO:0007411	axon guidance	*GLI3*
ENSG00000106571	GO:0000122	negative regulation of transcription from RNA polymerase II promoter	*GLI3*
ENSG00000106571	GO:0008285	negative regulation of cell proliferation	*GLI3*

ENSG00000106571	GO:0045893	positive regulation of transcription, DNA-dependent	*GLI3*
ENSG00000106571	GO:0043066	negative regulation of apoptotic process	*GLI3*
ENSG00000106571	GO:0042475	odontogenesis of dentin-containing tooth	*GLI3*
ENSG00000106571	GO:0045944	positive regulation of transcription from RNA polymerase II promoter	*GLI3*
ENSG00000106571	GO:0007507	heart development	*GLI3*
ENSG00000106571	GO:0048839	inner ear development	*GLI3*
ENSG00000106571	GO:0045892	negative regulation of transcription, DNA-dependent	*GLI3*
ENSG00000106571	GO:0009952	anterior/posterior pattern specification	*GLI3*
ENSG00000106571	GO:0030324	lung development	*GLI3*
ENSG00000106571	GO:0043586	tongue development	*GLI3*
ENSG00000106571	GO:0001701	in utero embryonic development	*GLI3*
ENSG00000106571	GO:0048589	developmental growth	*GLI3*
ENSG00000106571	GO:0002052	positive regulation of neuroblast proliferation	*GLI3*
ENSG00000106571	GO:0030318	melanocyte differentiation	*GLI3*
ENSG00000106571	GO:0048709	oligodendrocyte differentiation	*GLI3*
ENSG00000106571	GO:0090090	negative regulation of canonical Wnt receptor signaling pathway	*GLI3*
ENSG00000106571	GO:0045665	negative regulation of neuron differentiation	*GLI3*
ENSG00000106571	GO:0007224	smoothened signaling pathway	*GLI3*
ENSG00000106571	GO:0045669	positive regulation of osteoblast differentiation	*GLI3*
ENSG00000106571	GO:0045879	negative regulation of smoothened signaling pathway	*GLI3*
ENSG00000106571	GO:0001658	branching involved in ureteric bud morphogenesis	*GLI3*
ENSG00000106571	GO:0048704	embryonic skeletal system morphogenesis	*GLI3*
ENSG00000106571	GO:0021543	pallium development	*GLI3*
ENSG00000106571	GO:0021798	forebrain dorsal/ventral pattern formation	*GLI3*
ENSG00000106571	GO:0009954	proximal/distal pattern formation	*GLI3*
ENSG00000106571	GO:0042733	embryonic digit morphogenesis	*GLI3*
ENSG00000106571	GO:0021631	optic nerve morphogenesis	*GLI3*
ENSG00000106571	GO:0048593	camera-type eye morphogenesis	*GLI3*
ENSG00000106571	GO:0048557	embryonic digestive tract morphogenesis	*GLI3*
ENSG00000106571	GO:0060021	palate development	*GLI3*
ENSG00000106571	GO:0045060	negative thymic T cell selection	*GLI3*
ENSG00000106571	GO:0033077	T cell differentiation in thymus	*GLI3*
ENSG00000106571	GO:0001656	metanephros development	*GLI3*
ENSG00000106571	GO:0030879	mammary gland development	*GLI3*
ENSG00000106571	GO:0016485	protein processing	*GLI3*
ENSG00000106571	GO:0048663	neuron fate commitment	*GLI3*
ENSG00000106571	GO:0032332	positive regulation of chondrocyte differentiation	*GLI3*
ENSG00000106571	GO:0021915	neural tube development	*GLI3*
ENSG00000106571	GO:0042307	positive regulation of protein import into nucleus	*GLI3*
ENSG00000106571	GO:0021801	cerebral cortex radial glia guided migration	*GLI3*
ENSG00000106571	GO:0007442	hindgut morphogenesis	*GLI3*
ENSG00000106571	GO:0021513	spinal cord dorsal/ventral patterning	*GLI3*
ENSG00000106571	GO:0021544	subpallium development	*GLI3*
ENSG00000106571	GO:0021775	smoothened signaling pathway involved in ventral spinal cord interneuron specification	*GLI3*
ENSG00000106571	GO:0021776	smoothened signaling pathway involved in spinal cord motor neuron cell fate specification	*GLI3*

ENSG00000106571	GO:0021861	forebrain radial glial cell differentiation	*GLI3*
ENSG00000106571	GO:0022018	lateral ganglionic eminence cell proliferation	*GLI3*
ENSG00000106571	GO:0035295	tube development	*GLI3*
ENSG00000106571	GO:0046638	positive regulation of alpha-beta T cell differentiation	*GLI3*
ENSG00000106571	GO:0046639	negative regulation of alpha-beta T cell differentiation	*GLI3*
ENSG00000106571	GO:0048646	anatomical structure formation involved in morphogenesis	*GLI3*
ENSG00000106571	GO:0060364	frontal suture morphogenesis	*GLI3*
ENSG00000106571	GO:0060366	lambdoid suture morphogenesis	*GLI3*
ENSG00000106571	GO:0060367	sagittal suture morphogenesis	*GLI3*
ENSG00000106571	GO:0060594	mammary gland specification	*GLI3*
ENSG00000106571	GO:0060831	smoothened signaling pathway involved in dorsal/ventral neural tube patterning	*GLI3*
ENSG00000106571	GO:0060840	artery development	*GLI3*
ENSG00000106571	GO:0060873	anterior semicircular canal development	*GLI3*
ENSG00000106571	GO:0060875	lateral semicircular canal development	*GLI3*
ENSG00000106571	GO:0061005	cell differentiation involved in kidney development	*GLI3*
ENSG00000106571	GO:0070242	thymocyte apoptosis	*GLI3*
ENSG00000106571	GO:0035108	limb morphogenesis	*GLI3*
ENSG00000106571	GO:0048566	embryonic digestive tract development	*GLI3*
ENSG00000106571	GO:0043585	nose morphogenesis	*GLI3*
ENSG00000106571	GO:0060173	limb development	*GLI3*
ENSG00000106571	GO:0045595	regulation of cell differentiation	*GLI3*
ENSG00000106571	GO:0048754	branching morphogenesis of a tube	*GLI3*
ENSG00000106571	GO:0045596	negative regulation of cell differentiation	*GLI3*
ENSG00000106571	GO:0030326	embryonic limb morphogenesis	*GLI3*
ENSG00000106571	GO:0048856	anatomical structure development	*GLI3*
ENSG00000106571	GO:0048598	embryonic morphogenesis	*GLI3*
ENSG00000106571	GO:0021537	telencephalon development	*GLI3*
ENSG00000106571	GO:0030900	forebrain development	*GLI3*
ENSG00000106571	GO:0010468	regulation of gene expression	*GLI3*
ENSG00000106571	GO:0009953	dorsal/ventral pattern formation	*GLI3*
ENSG00000106571	GO:0007420	brain development	*GLI3*
ENSG00000106571	GO:0043010	camera-type eye development	*GLI3*
ENSG00000106571	GO:0006355	regulation of transcription, DNA-dependent	*GLI3*
ENSG00000106571	GO:0042981	regulation of apoptotic process	*GLI3*
ENSG00000106571	GO:0042127	regulation of cell proliferation	*GLI3*
ENSG00000106571	GO:0007417	central nervous system development	*GLI3*
ENSG00000106571	GO:0001822	kidney development	*GLI3*
ENSG00000106571	GO:0005634	nucleus	*GLI3*
ENSG00000106571	GO:0005737	cytoplasm	*GLI3*
ENSG00000106571	GO:0016592	mediator complex	*GLI3*
ENSG00000106571	GO:0005622	intracellular	*GLI3*
ENSG00000106571	GO:0005829	cytosol	*GLI3*
ENSG00000106571	GO:0005730	nucleolus	*GLI3*
ENSG00000106571	GO:0043231	intracellular membrane-bounded organelle	*GLI3*
ENSG00000106571	GO:0005929	cilium	*GLI3*
ENSG00000106571	GO:0017053	transcriptional repressor complex	*GLI3*
ENSG00000106571	GO:0003700	sequence-specific DNA binding transcription factor activity	*GLI3*
ENSG00000106571	GO:0008270	zinc ion binding	*GLI3*
ENSG00000106571	GO:0005515	protein binding	*GLI3*
ENSG00000106571	GO:0046872	metal ion binding	*GLI3*

ENSG00000106571	GO:0043565	sequence-specific DNA binding	*GLI3*
ENSG00000106571	GO:0042826	histone deacetylase binding	*GLI3*
ENSG00000106571	GO:0003682	chromatin binding	*GLI3*
ENSG00000106571	GO:0035035	histone acetyltransferase binding	*GLI3*
ENSG00000106571	GO:0008013	beta-catenin binding	*GLI3*
ENSG00000156049	GO:0007165	signal transduction	*GNA14*
ENSG00000156049	GO:0007186	G-protein coupled receptor signaling pathway	*GNA14*
ENSG00000156049	GO:0007188	G-protein signaling, coupled to cAMP nucleotide second messenger	*GNA14*
ENSG00000156049	GO:0006184	GTP catabolic process	*GNA14*
ENSG00000156049	GO:0030168	platelet activation	*GNA14*
ENSG00000156049	GO:0007596	blood coagulation	*GNA14*
ENSG00000156049	GO:0006471	protein ADP-ribosylation	*GNA14*
ENSG00000156049	GO:0060158	activation of phospholipase C activity by dopamine receptor signaling pathway	*GNA14*
ENSG00000156049	GO:0005886	plasma membrane	*GNA14*
ENSG00000156049	GO:0005834	heterotrimeric G-protein complex	*GNA14*
ENSG00000156049	GO:0031234	extrinsic to internal side of plasma membrane	*GNA14*
ENSG00000156049	GO:0000166	nucleotide binding	*GNA14*
ENSG00000156049	GO:0046872	metal ion binding	*GNA14*
ENSG00000156049	GO:0005525	GTP binding	*GNA14*
ENSG00000156049	GO:0005085	guanyl-nucleotide exchange factor activity	*GNA14*
ENSG00000156049	GO:0004871	signal transducer activity	*GNA14*
ENSG00000156049	GO:0019001	guanyl nucleotide binding	*GNA14*
ENSG00000156049	GO:0003924	GTPase activity	*GNA14*
ENSG00000156049	GO:0001664	G-protein coupled receptor binding	*GNA14*
ENSG00000156049	GO:0031683	G-protein beta/gamma-subunit complex binding	*GNA14*
ENSG00000156049	GO:0004629	phospholipase C activity	*GNA14*
ENSG00000112293	GO:0008285	negative regulation of cell proliferation	*GPLD1*
ENSG00000112293	GO:0002062	chondrocyte differentiation	*GPLD1*
ENSG00000112293	GO:0043065	positive regulation of apoptotic process	*GPLD1*
ENSG00000112293	GO:0071277	cellular response to calcium ion	*GPLD1*
ENSG00000112293	GO:0035774	positive regulation of insulin secretion involved in cellular response to glucose stimulus	*GPLD1*
ENSG00000112293	GO:0032869	cellular response to insulin stimulus	*GPLD1*
ENSG00000112293	GO:0008286	insulin receptor signaling pathway	*GPLD1*
ENSG00000112293	GO:0009749	response to glucose stimulus	*GPLD1*
ENSG00000112293	GO:0001503	ossification	*GPLD1*
ENSG00000112293	GO:0035690	cellular response to drug	*GPLD1*
ENSG00000112293	GO:0010867	positive regulation of triglyceride biosynthetic process	*GPLD1*
ENSG00000112293	GO:0010694	positive regulation of alkaline phosphatase activity	*GPLD1*
ENSG00000112293	GO:0002042	cell migration involved in sprouting angiogenesis	*GPLD1*
ENSG00000112293	GO:0071397	cellular response to cholesterol	*GPLD1*
ENSG00000112293	GO:0045919	positive regulation of cytolysis	*GPLD1*
ENSG00000112293	GO:0002430	complement receptor mediated signaling pathway	*GPLD1*
ENSG00000112293	GO:0046470	phosphatidylcholine metabolic process	*GPLD1*
ENSG00000112293	GO:0010595	positive regulation of endothelial cell migration	*GPLD1*
ENSG00000112293	GO:0051044	positive regulation of membrane protein	*GPLD1*

		ectodomain proteolysis	
ENSG00000112293	GO:0010897	negative regulation of triglyceride catabolic process	*GPLD1*
ENSG00000112293	GO:0010907	positive regulation of glucose metabolic process	*GPLD1*
ENSG00000112293	GO:0070633	transepithelial transport	*GPLD1*
ENSG00000112293	GO:0006507	GPI anchor release	*GPLD1*
ENSG00000112293	GO:0010983	positive regulation of high-density lipoprotein particle clearance	*GPLD1*
ENSG00000112293	GO:0035701	hematopoietic stem cell migration	*GPLD1*
ENSG00000112293	GO:0051047	positive regulation of secretion	*GPLD1*
ENSG00000112293	GO:0071401	cellular response to triglyceride	*GPLD1*
ENSG00000112293	GO:0071467	cellular response to pH	*GPLD1*
ENSG00000112293	GO:0097241	hematopoietic stem cell migration to bone marrow	*GPLD1*
ENSG00000112293	GO:1900076	regulation of cellular response to insulin stimulus	*GPLD1*
ENSG00000112293	GO:0005737	cytoplasm	*GPLD1*
ENSG00000112293	GO:0005576	extracellular region	*GPLD1*
ENSG00000112293	GO:0005622	intracellular	*GPLD1*
ENSG00000112293	GO:0005615	extracellular space	*GPLD1*
ENSG00000112293	GO:0005578	proteinaceous extracellular matrix	*GPLD1*
ENSG00000112293	GO:0005792	microsome	*GPLD1*
ENSG00000112293	GO:0044444	cytoplasmic part	*GPLD1*
ENSG00000112293	GO:0016788	hydrolase activity, acting on ester bonds	*GPLD1*
ENSG00000112293	GO:0004630	phospholipase D activity	*GPLD1*
ENSG00000112293	GO:0017080	sodium channel regulator activity	*GPLD1*
ENSG00000112293	GO:0004621	glycosylphosphatidylinositol phospholipase D activity	*GPLD1*
ENSG00000132975	GO:0007186	G-protein coupled receptor signaling pathway	*GPR12*
ENSG00000132975	GO:0006874	cellular calcium ion homeostasis	*GPR12*
ENSG00000132975	GO:0016021	integral to membrane	*GPR12*
ENSG00000132975	GO:0005886	plasma membrane	*GPR12*
ENSG00000132975	GO:0005887	integral to plasma membrane	*GPR12*
ENSG00000132975	GO:0004872	receptor activity	*GPR12*
ENSG00000132975	GO:0004871	signal transducer activity	*GPR12*
ENSG00000132975	GO:0004930	G-protein coupled receptor activity	*GPR12*
ENSG00000132975	GO:0031210	phosphatidylcholine binding	*GPR12*
ENSG00000151025	GO:0007186	G-protein coupled receptor signaling pathway	*GPR158*
ENSG00000151025	GO:0016021	integral to membrane	*GPR158*
ENSG00000151025	GO:0005886	plasma membrane	*GPR158*
ENSG00000151025	GO:0004872	receptor activity	*GPR158*
ENSG00000151025	GO:0004871	signal transducer activity	*GPR158*
ENSG00000151025	GO:0004930	G-protein coupled receptor activity	*GPR158*
ENSG00000177885	GO:0007411	axon guidance	*GRB2*
ENSG00000177885	GO:0030154	cell differentiation	*GRB2*
ENSG00000177885	GO:0051291	protein heterooligomerization	*GRB2*
ENSG00000177885	GO:0007568	aging	*GRB2*
ENSG00000177885	GO:0007265	Ras protein signal transduction	*GRB2*
ENSG00000177885	GO:0007267	cell-cell signaling	*GRB2*
ENSG00000177885	GO:0008286	insulin receptor signaling pathway	*GRB2*
ENSG00000177885	GO:0030168	platelet activation	*GRB2*
ENSG00000177885	GO:2000379	positive regulation of reactive oxygen species metabolic process	*GRB2*
ENSG00000177885	GO:0007596	blood coagulation	*GRB2*

ENSG00000177885	GO:0043408	regulation of MAPK cascade	*GRB2*
ENSG00000177885	GO:0050900	leukocyte migration	*GRB2*
ENSG00000177885	GO:0031623	receptor internalization	*GRB2*
ENSG00000177885	GO:0042770	signal transduction in response to DNA damage	*GRB2*
ENSG00000177885	GO:0007173	epidermal growth factor receptor signaling pathway	*GRB2*
ENSG00000177885	GO:0048646	anatomical structure formation involved in morphogenesis	*GRB2*
ENSG00000177885	GO:0071479	cellular response to ionizing radiation	*GRB2*
ENSG00000177885	GO:0031295	T cell costimulation	*GRB2*
ENSG00000177885	GO:0008543	fibroblast growth factor receptor signaling pathway	*GRB2*
ENSG00000177885	GO:0048011	nerve growth factor receptor signaling pathway	*GRB2*
ENSG00000177885	GO:0060670	branching involved in embryonic placenta morphogenesis	*GRB2*
ENSG00000177885	GO:0042059	negative regulation of epidermal growth factor receptor signaling pathway	*GRB2*
ENSG00000177885	GO:0044419	interspecies interaction between organisms	*GRB2*
ENSG00000177885	GO:0030838	positive regulation of actin filament polymerization	*GRB2*
ENSG00000177885	GO:0016020	membrane	*GRB2*
ENSG00000177885	GO:0005634	nucleus	*GRB2*
ENSG00000177885	GO:0005737	cytoplasm	*GRB2*
ENSG00000177885	GO:0005829	cytosol	*GRB2*
ENSG00000177885	GO:0005794	Golgi apparatus	*GRB2*
ENSG00000177885	GO:0005768	endosome	*GRB2*
ENSG00000177885	GO:0043234	protein complex	*GRB2*
ENSG00000177885	GO:0012506	vesicle membrane	*GRB2*
ENSG00000177885	GO:0005886	plasma membrane	*GRB2*
ENSG00000177885	GO:0005515	protein binding	*GRB2*
ENSG00000177885	GO:0005154	epidermal growth factor receptor binding	*GRB2*
ENSG00000177885	GO:0019904	protein domain specific binding	*GRB2*
ENSG00000177885	GO:0019903	protein phosphatase binding	*GRB2*
ENSG00000177885	GO:0005070	SH3/SH2 adaptor activity	*GRB2*
ENSG00000177885	GO:0001784	phosphotyrosine binding	*GRB2*
ENSG00000177885	GO:0046875	ephrin receptor binding	*GRB2*
ENSG00000177885	GO:0051219	phosphoprotein binding	*GRB2*
ENSG00000177885	GO:0043560	insulin receptor substrate binding	*GRB2*
ENSG00000177885	GO:0005168	neurotrophin TRKA receptor binding	*GRB2*
ENSG00000163041	GO:0006334	nucleosome assembly	*H3F3A*
ENSG00000163041	GO:0007596	blood coagulation	*H3F3A*
ENSG00000163041	GO:0005634	nucleus	*H3F3A*
ENSG00000163041	GO:0005576	extracellular region	*H3F3A*
ENSG00000163041	GO:0005694	chromosome	*H3F3A*
ENSG00000163041	GO:0000786	nucleosome	*H3F3A*
ENSG00000163041	GO:0005654	nucleoplasm	*H3F3A*
ENSG00000163041	GO:0003677	DNA binding	*H3F3A*
ENSG00000132475	GO:0006334	nucleosome assembly	*H3F3B*
ENSG00000132475	GO:0007596	blood coagulation	*H3F3B*
ENSG00000132475	GO:0005634	nucleus	*H3F3B*
ENSG00000132475	GO:0005576	extracellular region	*H3F3B*
ENSG00000132475	GO:0005694	chromosome	*H3F3B*
ENSG00000132475	GO:0000786	nucleosome	*H3F3B*
ENSG00000132475	GO:0005654	nucleoplasm	*H3F3B*

ENSG00000132475	GO:0003677	DNA binding	*H3F3B*
ENSG00000101323	GO:0006979	response to oxidative stress	*HAO1*
ENSG00000101323	GO:0055114	oxidation-reduction process	*HAO1*
ENSG00000101323	GO:0044281	small molecule metabolic process	*HAO1*
ENSG00000101323	GO:0034641	cellular nitrogen compound metabolic process	*HAO1*
ENSG00000101323	GO:0006537	glutamate biosynthetic process	*HAO1*
ENSG00000101323	GO:0001561	fatty acid alpha-oxidation	*HAO1*
ENSG00000101323	GO:0046487	glyoxylate metabolic process	*HAO1*
ENSG00000101323	GO:0046296	glycolate catabolic process	*HAO1*
ENSG00000101323	GO:0005777	peroxisome	*HAO1*
ENSG00000101323	GO:0005782	peroxisomal matrix	*HAO1*
ENSG00000101323	GO:0016491	oxidoreductase activity	*HAO1*
ENSG00000101323	GO:0010181	FMN binding	*HAO1*
ENSG00000101323	GO:0016638	oxidoreductase activity, acting on the CH-NH2 group of donors	*HAO1*
ENSG00000101323	GO:0003973	(S)-2-hydroxy-acid oxidase activity	*HAO1*
ENSG00000101323	GO:0052852	very-long-chain-(S)-2-hydroxy-acid oxidase activity	*HAO1*
ENSG00000101323	GO:0052853	long-chain-(S)-2-hydroxy-long-chain-acid oxidase activity	*HAO1*
ENSG00000101323	GO:0052854	medium-chain-(S)-2-hydroxy-acid oxidase activity	*HAO1*
ENSG00000101323	GO:0008891	glycolate oxidase activity	*HAO1*
ENSG00000101323	GO:0047969	glyoxylate oxidase activity	*HAO1*
ENSG00000101323	GO:0015930	glutamate synthase activity	*HAO1*
ENSG00000119285	GO:0006364	rRNA processing	*HEATR1*
ENSG00000119285	GO:0005634	nucleus	*HEATR1*
ENSG00000119285	GO:0005730	nucleolus	*HEATR1*
ENSG00000119285	GO:0030529	ribonucleoprotein complex	*HEATR1*
ENSG00000119285	GO:0005488	binding	*HEATR1*
ENSG00000176387	GO:0008152	metabolic process	*HSD11B2*
ENSG00000176387	GO:0042493	response to drug	*HSD11B2*
ENSG00000176387	GO:0032868	response to insulin stimulus	*HSD11B2*
ENSG00000176387	GO:0007565	female pregnancy	*HSD11B2*
ENSG00000176387	GO:0001666	response to hypoxia	*HSD11B2*
ENSG00000176387	GO:0051384	response to glucocorticoid stimulus	*HSD11B2*
ENSG00000176387	GO:0032094	response to food	*HSD11B2*
ENSG00000176387	GO:0048545	response to steroid hormone stimulus	*HSD11B2*
ENSG00000176387	GO:0006704	glucocorticoid biosynthetic process	*HSD11B2*
ENSG00000176387	GO:0002017	regulation of blood volume by renal aldosterone	*HSD11B2*
ENSG00000176387	GO:0055114	oxidation-reduction process	*HSD11B2*
ENSG00000176387	GO:0006950	response to stress	*HSD11B2*
ENSG00000176387	GO:0008211	glucocorticoid metabolic process	*HSD11B2*
ENSG00000176387	GO:0005737	cytoplasm	*HSD11B2*
ENSG00000176387	GO:0005783	endoplasmic reticulum	*HSD11B2*
ENSG00000176387	GO:0005792	microsome	*HSD11B2*
ENSG00000176387	GO:0016491	oxidoreductase activity	*HSD11B2*
ENSG00000176387	GO:0051287	NAD binding	*HSD11B2*
ENSG00000176387	GO:0005496	steroid binding	*HSD11B2*
ENSG00000176387	GO:0003845	11-beta-hydroxysteroid dehydrogenase [NAD(P)] activity	*HSD11B2*
ENSG00000134049	GO:0042981	regulation of apoptotic process	*IER3IP1*
ENSG00000134049	GO:0016020	membrane	*IER3IP1*
ENSG00000134049	GO:0016021	integral to membrane	*IER3IP1*

ENSG00000134049	GO:0005783	endoplasmic reticulum	*IER3IP1*
ENSG00000134049	GO:0005789	endoplasmic reticulum membrane	*IER3IP1*
ENSG00000134049	GO:0005794	Golgi apparatus	*IER3IP1*
ENSG00000027697	GO:0007165	signal transduction	*IFNGR1*
ENSG00000027697	GO:0009615	response to virus	*IFNGR1*
ENSG00000027697	GO:0019221	cytokine-mediated signaling pathway	*IFNGR1*
ENSG00000027697	GO:0060333	interferon-gamma-mediated signaling pathway	*IFNGR1*
ENSG00000027697	GO:0060334	regulation of interferon-gamma-mediated signaling pathway	*IFNGR1*
ENSG00000027697	GO:0016020	membrane	*IFNGR1*
ENSG00000027697	GO:0005886	plasma membrane	*IFNGR1*
ENSG00000027697	GO:0005887	integral to plasma membrane	*IFNGR1*
ENSG00000027697	GO:0004872	receptor activity	*IFNGR1*
ENSG00000027697	GO:0019955	cytokine binding	*IFNGR1*
ENSG00000027697	GO:0004906	interferon-gamma receptor activity	*IFNGR1*
ENSG00000115590	GO:0006955	immune response	*IL1R2*
ENSG00000115590	GO:0016021	integral to membrane	*IL1R2*
ENSG00000115590	GO:0005886	plasma membrane	*IL1R2*
ENSG00000115590	GO:0005576	extracellular region	*IL1R2*
ENSG00000115590	GO:0004872	receptor activity	*IL1R2*
ENSG00000115590	GO:0005515	protein binding	*IL1R2*
ENSG00000115590	GO:0004908	interleukin-1 receptor activity	*IL1R2*
ENSG00000115590	GO:0004910	interleukin-1, Type II, blocking receptor activity	*IL1R2*
ENSG00000164675	GO:0001669	acrosomal vesicle	*IQUB*
ENSG00000164675	GO:0009434	microtubule-based flagellum	*IQUB*
ENSG00000164675	GO:0005515	protein binding	*IQUB*
ENSG00000140968	GO:0000122	negative regulation of transcription from RNA polymerase II promoter	*IRF8*
ENSG00000140968	GO:0045893	positive regulation of transcription, DNA-dependent	*IRF8*
ENSG00000140968	GO:0006955	immune response	*IRF8*
ENSG00000140968	GO:0006355	regulation of transcription, DNA-dependent	*IRF8*
ENSG00000140968	GO:0042742	defense response to bacterium	*IRF8*
ENSG00000140968	GO:0071222	cellular response to lipopolysaccharide	*IRF8*
ENSG00000140968	GO:0019221	cytokine-mediated signaling pathway	*IRF8*
ENSG00000140968	GO:0032729	positive regulation of interferon-gamma production	*IRF8*
ENSG00000140968	GO:0044130	negative regulation of growth of symbiont in host	*IRF8*
ENSG00000140968	GO:0006909	phagocytosis	*IRF8*
ENSG00000140968	GO:0042832	defense response to protozoan	*IRF8*
ENSG00000140968	GO:0032735	positive regulation of interleukin-12 production	*IRF8*
ENSG00000140968	GO:0030099	myeloid cell differentiation	*IRF8*
ENSG00000140968	GO:0060333	interferon-gamma-mediated signaling pathway	*IRF8*
ENSG00000140968	GO:0060337	type I interferon-mediated signaling pathway	*IRF8*
ENSG00000140968	GO:0009617	response to bacterium	*IRF8*
ENSG00000140968	GO:0005634	nucleus	*IRF8*
ENSG00000140968	GO:0005829	cytosol	*IRF8*
ENSG00000140968	GO:0003700	sequence-specific DNA binding transcription factor activity	*IRF8*
ENSG00000140968	GO:0003705	sequence-specific distal enhancer binding RNA polymerase II transcription factor activity	*IRF8*

ENSG00000140968	GO:0000975	regulatory region DNA binding	*IRF8*
ENSG00000140968	GO:0003677	DNA binding	*IRF8*
ENSG00000140968	GO:0005515	protein binding	*IRF8*
ENSG00000175329	GO:0006355	regulation of transcription, DNA-dependent	*ISX*
ENSG00000175329	GO:0005634	nucleus	*ISX*
ENSG00000175329	GO:0003700	sequence-specific DNA binding transcription factor activity	*ISX*
ENSG00000175329	GO:0043565	sequence-specific DNA binding	*ISX*
ENSG00000144668	GO:0007155	cell adhesion	*ITGA9*
ENSG00000144668	GO:0007229	integrin-mediated signaling pathway	*ITGA9*
ENSG00000144668	GO:0016020	membrane	*ITGA9*
ENSG00000144668	GO:0016021	integral to membrane	*ITGA9*
ENSG00000144668	GO:0008305	integrin complex	*ITGA9*
ENSG00000144668	GO:0004872	receptor activity	*ITGA9*
ENSG00000144668	GO:0007411	axon guidance	*ITGA9*
ENSG00000144668	GO:0042060	wound healing	*ITGA9*
ENSG00000144668	GO:0030593	neutrophil chemotaxis	*ITGA9*
ENSG00000144668	GO:0005886	plasma membrane	*ITGA9*
ENSG00000144668	GO:0005624	membrane fraction	*ITGA9*
ENSG00000144668	GO:0034679	alpha9-beta1 integrin complex	*ITGA9*
ENSG00000144668	GO:0009925	basal plasma membrane	*ITGA9*
ENSG00000144668	GO:0043236	laminin binding	*ITGA9*
ENSG00000144668	GO:0005518	collagen binding	*ITGA9*
ENSG00000144668	GO:0005515	protein binding	*ITGA9*
ENSG00000113263	GO:0035556	intracellular signal transduction	*ITK*
ENSG00000113263	GO:0050852	T cell receptor signaling pathway	*ITK*
ENSG00000113263	GO:0007202	activation of phospholipase C activity	*ITK*
ENSG00000113263	GO:0001865	NK T cell differentiation	*ITK*
ENSG00000113263	GO:0002250	adaptive immune response	*ITK*
ENSG00000113263	GO:0032609	interferon-gamma production	*ITK*
ENSG00000113263	GO:0032633	interleukin-4 production	*ITK*
ENSG00000113263	GO:0005829	cytosol	*ITK*
ENSG00000113263	GO:0007165	signal transduction	*ITK*
ENSG00000113263	GO:0001816	cytokine production	*ITK*
ENSG00000113263	GO:0006468	protein phosphorylation	*ITK*
ENSG00000113263	GO:0006968	cellular defense response	*ITK*
ENSG00000113263	GO:0042110	T cell activation	*ITK*
ENSG00000113263	GO:0005737	cytoplasm	*ITK*
ENSG00000113263	GO:0000166	nucleotide binding	*ITK*
ENSG00000113263	GO:0005515	protein binding	*ITK*
ENSG00000113263	GO:0005524	ATP binding	*ITK*
ENSG00000113263	GO:0005543	phospholipid binding	*ITK*
ENSG00000113263	GO:0046872	metal ion binding	*ITK*
ENSG00000113263	GO:0004672	protein kinase activity	*ITK*
ENSG00000113263	GO:0004713	protein tyrosine kinase activity	*ITK*
ENSG00000113263	GO:0004715	non-membrane spanning protein tyrosine kinase activity	*ITK*
ENSG00000184408	GO:0006811	ion transport	*KCND2*
ENSG00000184408	GO:0006813	potassium ion transport	*KCND2*
ENSG00000184408	GO:0055085	transmembrane transport	*KCND2*
ENSG00000184408	GO:0051291	protein heterooligomerization	*KCND2*
ENSG00000184408	GO:0007268	synaptic transmission	*KCND2*
ENSG00000184408	GO:0001508	regulation of action potential	*KCND2*
ENSG00000184408	GO:0071805	potassium ion transmembrane transport	*KCND2*
ENSG00000184408	GO:0016020	membrane	*KCND2*

ENSG00000184408	GO:0005886	plasma membrane	*KCND2*
ENSG00000184408	GO:0042995	cell projection	*KCND2*
ENSG00000184408	GO:0008076	voltage-gated potassium channel complex	*KCND2*
ENSG00000184408	GO:0009986	cell surface	*KCND2*
ENSG00000184408	GO:0043025	neuronal cell body	*KCND2*
ENSG00000184408	GO:0019717	synaptosome	*KCND2*
ENSG00000184408	GO:0014069	postsynaptic density	*KCND2*
ENSG00000184408	GO:0043197	dendritic spine	*KCND2*
ENSG00000184408	GO:0030425	dendrite	*KCND2*
ENSG00000184408	GO:0005887	integral to plasma membrane	*KCND2*
ENSG00000184408	GO:0005216	ion channel activity	*KCND2*
ENSG00000184408	GO:0005515	protein binding	*KCND2*
ENSG00000184408	GO:0005249	voltage-gated potassium channel activity	*KCND2*
ENSG00000184408	GO:0046872	metal ion binding	*KCND2*
ENSG00000184408	GO:0046982	protein heterodimerization activity	*KCND2*
ENSG00000184408	GO:0005244	voltage-gated ion channel activity	*KCND2*
ENSG00000184408	GO:0005250	A-type (transient outward) potassium channel activity	*KCND2*
ENSG00000184408	GO:0005267	potassium channel activity	*KCND2*
ENSG00000135750	GO:0006811	ion transport	*KCNK1*
ENSG00000135750	GO:0006813	potassium ion transport	*KCNK1*
ENSG00000135750	GO:0035094	response to nicotine	*KCNK1*
ENSG00000135750	GO:0007268	synaptic transmission	*KCNK1*
ENSG00000135750	GO:0071805	potassium ion transmembrane transport	*KCNK1*
ENSG00000135750	GO:0016020	membrane	*KCNK1*
ENSG00000135750	GO:0016021	integral to membrane	*KCNK1*
ENSG00000135750	GO:0005737	cytoplasm	*KCNK1*
ENSG00000135750	GO:0005886	plasma membrane	*KCNK1*
ENSG00000135750	GO:0008076	voltage-gated potassium channel complex	*KCNK1*
ENSG00000135750	GO:0005768	endosome	*KCNK1*
ENSG00000135750	GO:0031526	brush border membrane	*KCNK1*
ENSG00000135750	GO:0016324	apical plasma membrane	*KCNK1*
ENSG00000135750	GO:0005242	inward rectifier potassium channel activity	*KCNK1*
ENSG00000135750	GO:0005244	voltage-gated ion channel activity	*KCNK1*
ENSG00000135750	GO:0005267	potassium channel activity	*KCNK1*
ENSG00000135750	GO:0005216	ion channel activity	*KCNK1*
ENSG00000156113	GO:0006811	ion transport	*KCNMA1*
ENSG00000156113	GO:0006813	potassium ion transport	*KCNMA1*
ENSG00000156113	GO:0055085	transmembrane transport	*KCNMA1*
ENSG00000156113	GO:0016020	membrane	*KCNMA1*
ENSG00000156113	GO:0016021	integral to membrane	*KCNMA1*
ENSG00000156113	GO:0008076	voltage-gated potassium channel complex	*KCNMA1*
ENSG00000156113	GO:0000166	nucleotide binding	*KCNMA1*
ENSG00000156113	GO:0005216	ion channel activity	*KCNMA1*
ENSG00000156113	GO:0005249	voltage-gated potassium channel activity	*KCNMA1*
ENSG00000156113	GO:0060072	large conductance calcium-activated potassium channel activity	*KCNMA1*
ENSG00000156113	GO:0015269	calcium-activated potassium channel activity	*KCNMA1*
ENSG00000156113	GO:0043065	positive regulation of apoptotic process	*KCNMA1*
ENSG00000156113	GO:0007268	synaptic transmission	*KCNMA1*
ENSG00000156113	GO:0006970	response to osmotic stress	*KCNMA1*
ENSG00000156113	GO:0071805	potassium ion transmembrane transport	*KCNMA1*
ENSG00000156113	GO:0001666	response to hypoxia	*KCNMA1*
ENSG00000156113	GO:0007596	blood coagulation	*KCNMA1*
ENSG00000156113	GO:0042391	regulation of membrane potential	*KCNMA1*

ENSG00000156113	GO:0051592	response to calcium ion	*KCNMA1*
ENSG00000156113	GO:0060083	smooth muscle contraction involved in micturition	*KCNMA1*
ENSG00000156113	GO:0045794	negative regulation of cell volume	*KCNMA1*
ENSG00000156113	GO:0034465	response to carbon monoxide	*KCNMA1*
ENSG00000156113	GO:0030007	cellular potassium ion homeostasis	*KCNMA1*
ENSG00000156113	GO:0060073	micturition	*KCNMA1*
ENSG00000156113	GO:0005886	plasma membrane	*KCNMA1*
ENSG00000156113	GO:0016324	apical plasma membrane	*KCNMA1*
ENSG00000156113	GO:0005901	caveola	*KCNMA1*
ENSG00000156113	GO:0005515	protein binding	*KCNMA1*
ENSG00000156113	GO:0046872	metal ion binding	*KCNMA1*
ENSG00000156113	GO:0003779	actin binding	*KCNMA1*
ENSG00000156113	GO:0005244	voltage-gated ion channel activity	*KCNMA1*
ENSG00000156113	GO:0005267	potassium channel activity	*KCNMA1*
ENSG00000156113	GO:0060087	relaxation of vascular smooth muscle	*KCNMA1*
ENSG00000156113	GO:0060082	eye blink reflex	*KCNMA1*
ENSG00000156113	GO:0032344	regulation of aldosterone metabolic process	*KCNMA1*
ENSG00000156113	GO:0042311	vasodilation	*KCNMA1*
ENSG00000156113	GO:0042491	auditory receptor cell differentiation	*KCNMA1*
ENSG00000156113	GO:0046541	saliva secretion	*KCNMA1*
ENSG00000156113	GO:0045475	locomotor rhythm	*KCNMA1*
ENSG00000156113	GO:0019228	regulation of action potential in neuron	*KCNMA1*
ENSG00000156113	GO:0048469	cell maturation	*KCNMA1*
ENSG00000156113	GO:0007623	circadian rhythm	*KCNMA1*
ENSG00000156113	GO:0050885	neuromuscular process controlling balance	*KCNMA1*
ENSG00000156113	GO:0007605	sensory perception of sound	*KCNMA1*
ENSG00000156113	GO:0007628	adult walking behavior	*KCNMA1*
ENSG00000156113	GO:0051260	protein homooligomerization	*KCNMA1*
ENSG00000156113	GO:0044444	cytoplasmic part	*KCNMA1*
ENSG00000156113	GO:0043195	terminal button	*KCNMA1*
ENSG00000156113	GO:0045211	postsynaptic membrane	*KCNMA1*
ENSG00000156113	GO:0009897	external side of plasma membrane	*KCNMA1*
ENSG00000156113	GO:0005783	endoplasmic reticulum	*KCNMA1*
ENSG00000197993	GO:0006508	proteolysis	*KEL*
ENSG00000197993	GO:0042310	vasoconstriction	*KEL*
ENSG00000197993	GO:0016021	integral to membrane	*KEL*
ENSG00000197993	GO:0005886	plasma membrane	*KEL*
ENSG00000197993	GO:0004222	metalloendopeptidase activity	*KEL*
ENSG00000197993	GO:0005515	protein binding	*KEL*
ENSG00000197993	GO:0046872	metal ion binding	*KEL*
ENSG00000197993	GO:0008233	peptidase activity	*KEL*
ENSG00000197993	GO:0008237	metallopeptidase activity	*KEL*
ENSG00000260040	GO:0006508	proteolysis	*KEL*
ENSG00000260040	GO:0042310	vasoconstriction	*KEL*
ENSG00000260040	GO:0016021	integral to membrane	*KEL*
ENSG00000260040	GO:0005886	plasma membrane	*KEL*
ENSG00000260040	GO:0004222	metalloendopeptidase activity	*KEL*
ENSG00000260040	GO:0005515	protein binding	*KEL*
ENSG00000260040	GO:0046872	metal ion binding	*KEL*
ENSG00000260040	GO:0008233	peptidase activity	*KEL*
ENSG00000260040	GO:0008237	metallopeptidase activity	*KEL*
ENSG00000149571	GO:0030097	hemopoiesis	*KIRREL3*
ENSG00000149571	GO:0016020	membrane	*KIRREL3*
ENSG00000149571	GO:0016021	integral to membrane	*KIRREL3*

ENSG00000149571	GO:0005886	plasma membrane	*KIRREL3*
ENSG00000149571	GO:0005576	extracellular region	*KIRREL3*
ENSG00000149571	GO:0030424	axon	*KIRREL3*
ENSG00000149571	GO:0043198	dendritic shaft	*KIRREL3*
ENSG00000149571	GO:0005515	protein binding	*KIRREL3*
ENSG00000049130	GO:0007155	cell adhesion	*KITLG*
ENSG00000049130	GO:0016020	membrane	*KITLG*
ENSG00000049130	GO:0005173	stem cell factor receptor binding	*KITLG*
ENSG00000049130	GO:0008283	cell proliferation	*KITLG*
ENSG00000049130	GO:0008284	positive regulation of cell proliferation	*KITLG*
ENSG00000049130	GO:0007165	signal transduction	*KITLG*
ENSG00000049130	GO:0043066	negative regulation of apoptotic process	*KITLG*
ENSG00000049130	GO:0008584	male gonad development	*KITLG*
ENSG00000049130	GO:0030097	hemopoiesis	*KITLG*
ENSG00000049130	GO:0001541	ovarian follicle development	*KITLG*
ENSG00000049130	GO:0001755	neural crest cell migration	*KITLG*
ENSG00000049130	GO:0050731	positive regulation of peptidyl-tyrosine phosphorylation	*KITLG*
ENSG00000049130	GO:0045740	positive regulation of DNA replication	*KITLG*
ENSG00000049130	GO:0035162	embryonic hemopoiesis	*KITLG*
ENSG00000049130	GO:0033026	negative regulation of mast cell apoptosis	*KITLG*
ENSG00000049130	GO:0070668	positive regulation of mast cell proliferation	*KITLG*
ENSG00000049130	GO:0046579	positive regulation of Ras protein signal transduction	*KITLG*
ENSG00000049130	GO:0035234	germ cell programmed cell death	*KITLG*
ENSG00000049130	GO:0045636	positive regulation of melanocyte differentiation	*KITLG*
ENSG00000049130	GO:0002763	positive regulation of myeloid leukocyte differentiation	*KITLG*
ENSG00000049130	GO:0016021	integral to membrane	*KITLG*
ENSG00000049130	GO:0005737	cytoplasm	*KITLG*
ENSG00000049130	GO:0005886	plasma membrane	*KITLG*
ENSG00000049130	GO:0005576	extracellular region	*KITLG*
ENSG00000049130	GO:0005615	extracellular space	*KITLG*
ENSG00000049130	GO:0005856	cytoskeleton	*KITLG*
ENSG00000049130	GO:0005515	protein binding	*KITLG*
ENSG00000049130	GO:0008083	growth factor activity	*KITLG*
ENSG00000049130	GO:0005125	cytokine activity	*KITLG*
ENSG00000116977	GO:0005529	sugar binding	*LGALS8*
ENSG00000116977	GO:0005737	cytoplasm	*LGALS8*
ENSG00000116977	GO:0005615	extracellular space	*LGALS8*
ENSG00000072201	GO:0005515	protein binding	*LNX1*
ENSG00000072201	GO:0051260	protein homooligomerization	*LNX1*
ENSG00000072201	GO:0042787	protein ubiquitination involved in ubiquitin-dependent protein catabolic process	*LNX1*
ENSG00000072201	GO:0006511	ubiquitin-dependent protein catabolic process	*LNX1*
ENSG00000072201	GO:0005737	cytoplasm	*LNX1*
ENSG00000072201	GO:0008270	zinc ion binding	*LNX1*
ENSG00000072201	GO:0016874	ligase activity	*LNX1*
ENSG00000072201	GO:0046872	metal ion binding	*LNX1*
ENSG00000072201	GO:0004842	ubiquitin-protein ligase activity	*LNX1*
ENSG00000072201	GO:0030165	PDZ domain binding	*LNX1*
ENSG00000081479	GO:0008283	cell proliferation	*LRP2*
ENSG00000081479	GO:0044281	small molecule metabolic process	*LRP2*
ENSG00000081479	GO:0031100	organ regeneration	*LRP2*

ENSG00000081479	GO:0006629	lipid metabolic process	*LRP2*
ENSG00000081479	GO:0016197	endosome transport	*LRP2*
ENSG00000081479	GO:0042493	response to drug	*LRP2*
ENSG00000081479	GO:0006486	protein glycosylation	*LRP2*
ENSG00000081479	GO:0007568	aging	*LRP2*
ENSG00000081479	GO:0006897	endocytosis	*LRP2*
ENSG00000081479	GO:0032526	response to retinoic acid	*LRP2*
ENSG00000081479	GO:0006898	receptor-mediated endocytosis	*LRP2*
ENSG00000081479	GO:0030900	forebrain development	*LRP2*
ENSG00000081479	GO:0010165	response to X-ray	*LRP2*
ENSG00000081479	GO:0033280	response to vitamin D	*LRP2*
ENSG00000081479	GO:0042359	vitamin D metabolic process	*LRP2*
ENSG00000081479	GO:0045056	transcytosis	*LRP2*
ENSG00000081479	GO:0046879	hormone secretion	*LRP2*
ENSG00000081479	GO:0008202	steroid metabolic process	*LRP2*
ENSG00000081479	GO:0006766	vitamin metabolic process	*LRP2*
ENSG00000081479	GO:0020028	hemoglobin import	*LRP2*
ENSG00000081479	GO:0042953	lipoprotein transport	*LRP2*
ENSG00000081479	GO:0016021	integral to membrane	*LRP2*
ENSG00000081479	GO:0005737	cytoplasm	*LRP2*
ENSG00000081479	GO:0005886	plasma membrane	*LRP2*
ENSG00000081479	GO:0005783	endoplasmic reticulum	*LRP2*
ENSG00000081479	GO:0005624	membrane fraction	*LRP2*
ENSG00000081479	GO:0005794	Golgi apparatus	*LRP2*
ENSG00000081479	GO:0005615	extracellular space	*LRP2*
ENSG00000081479	GO:0005764	lysosome	*LRP2*
ENSG00000081479	GO:0005768	endosome	*LRP2*
ENSG00000081479	GO:0031526	brush border membrane	*LRP2*
ENSG00000081479	GO:0016324	apical plasma membrane	*LRP2*
ENSG00000081479	GO:0005905	coated pit	*LRP2*
ENSG00000081479	GO:0030139	endocytic vesicle	*LRP2*
ENSG00000081479	GO:0005903	brush border	*LRP2*
ENSG00000081479	GO:0045177	apical part of cell	*LRP2*
ENSG00000081479	GO:0004872	receptor activity	*LRP2*
ENSG00000081479	GO:0005515	protein binding	*LRP2*
ENSG00000081479	GO:0005509	calcium ion binding	*LRP2*
ENSG00000081479	GO:0017124	SH3 domain binding	*LRP2*
ENSG00000081479	GO:0030492	hemoglobin binding	*LRP2*
ENSG00000081479	GO:0042954	lipoprotein transporter activity	*LRP2*
ENSG00000081479	GO:0016020	membrane	*LRP2*
ENSG00000183742	GO:0051781	positive regulation of cell division	*MACC1*
ENSG00000183742	GO:0006355	regulation of transcription, DNA-dependent	*MACC1*
ENSG00000183742	GO:0005634	nucleus	*MACC1*
ENSG00000183742	GO:0005737	cytoplasm	*MACC1*
ENSG00000183742	GO:0008083	growth factor activity	*MACC1*
ENSG00000172264	GO:0007420	brain development	*MACROD2*
ENSG00000172264	GO:0042278	purine nucleoside metabolic process	*MACROD2*
ENSG00000172264	GO:0005575	cellular_component	*MACROD2*
ENSG00000172264	GO:0019213	deacetylase activity	*MACROD2*
ENSG00000155130	GO:0044281	small molecule metabolic process	*MARCKS*
ENSG00000155130	GO:0050796	regulation of insulin secretion	*MARCKS*
ENSG00000155130	GO:0006112	energy reserve metabolic process	*MARCKS*
ENSG00000155130	GO:0005737	cytoplasm	*MARCKS*
ENSG00000155130	GO:0005886	plasma membrane	*MARCKS*
ENSG00000155130	GO:0005938	cell cortex	*MARCKS*

ENSG00000155130	GO:0005813	centrosome	*MARCKS*
ENSG00000155130	GO:0015629	actin cytoskeleton	*MARCKS*
ENSG00000155130	GO:0042585	germinal vesicle	*MARCKS*
ENSG00000155130	GO:0016020	membrane	*MARCKS*
ENSG00000155130	GO:0005516	calmodulin binding	*MARCKS*
ENSG00000155130	GO:0051015	actin filament binding	*MARCKS*
ENSG00000155130	GO:0005080	protein kinase C binding	*MARCKS*
ENSG00000155130	GO:0005515	protein binding	*MARCKS*
ENSG00000171444	GO:0016055	Wnt receptor signaling pathway	*MCC*
ENSG00000171444	GO:0007165	signal transduction	*MCC*
ENSG00000171444	GO:0090090	negative regulation of canonical Wnt receptor signaling pathway	*MCC*
ENSG00000171444	GO:0050680	negative regulation of epithelial cell proliferation	*MCC*
ENSG00000171444	GO:0010633	negative regulation of epithelial cell migration	*MCC*
ENSG00000171444	GO:0005634	nucleus	*MCC*
ENSG00000171444	GO:0005737	cytoplasm	*MCC*
ENSG00000171444	GO:0005886	plasma membrane	*MCC*
ENSG00000171444	GO:0004872	receptor activity	*MCC*
ENSG00000171444	GO:0005515	protein binding	*MCC*
ENSG00000171444	GO:0005509	calcium ion binding	*MCC*
ENSG00000145794	GO:0007155	cell adhesion	*MEGF10*
ENSG00000145794	GO:0043654	recognition of apoptotic cell	*MEGF10*
ENSG00000145794	GO:0006909	phagocytosis	*MEGF10*
ENSG00000145794	GO:0016021	integral to membrane	*MEGF10*
ENSG00000145794	GO:0005886	plasma membrane	*MEGF10*
ENSG00000145794	GO:0042995	cell projection	*MEGF10*
ENSG00000145794	GO:0001891	phagocytic cup	*MEGF10*
ENSG00000145794	GO:0016323	basolateral plasma membrane	*MEGF10*
ENSG00000145794	GO:0005515	protein binding	*MEGF10*
ENSG00000133808	GO:0005737	cytoplasm	*MICALCL*
ENSG00000133808	GO:0051019	mitogen-activated protein kinase binding	*MICALCL*
ENSG00000133808	GO:0008150	biological_process	*MICALCL*
ENSG00000133808	GO:0007275	multicellular organismal development	*MICALCL*
ENSG00000133808	GO:0030154	cell differentiation	*MICALCL*
ENSG00000133808	GO:0007283	spermatogenesis	*MICALCL*
ENSG00000182208	GO:0005634	nucleus	*MOB2*
ENSG00000182208	GO:0005737	cytoplasm	*MOB2*
ENSG00000182208	GO:0048471	perinuclear region of cytoplasm	*MOB2*
ENSG00000182208	GO:0005730	nucleolus	*MOB2*
ENSG00000182208	GO:0046872	metal ion binding	*MOB2*
ENSG00000129422	GO:0005634	nucleus	*MTUS1*
ENSG00000129422	GO:0005737	cytoplasm	*MTUS1*
ENSG00000129422	GO:0005886	plasma membrane	*MTUS1*
ENSG00000129422	GO:0005739	mitochondrion	*MTUS1*
ENSG00000129422	GO:0005794	Golgi apparatus	*MTUS1*
ENSG00000129422	GO:0005856	cytoskeleton	*MTUS1*
ENSG00000129422	GO:0005819	spindle	*MTUS1*
ENSG00000129422	GO:0005874	microtubule	*MTUS1*
ENSG00000129422	GO:0005815	microtubule organizing center	*MTUS1*
ENSG00000172915	GO:0008104	protein localization	*NBEA*
ENSG00000172915	GO:0005886	plasma membrane	*NBEA*
ENSG00000172915	GO:0005829	cytosol	*NBEA*
ENSG00000172915	GO:0005802	trans-Golgi network	*NBEA*
ENSG00000172915	GO:0012505	endomembrane system	*NBEA*

ENSG00000172915	GO:0005488	binding	*NBEA*
ENSG00000172915	GO:0005515	protein binding	*NBEA*
ENSG00000172915	GO:0019901	protein kinase binding	*NBEA*
ENSG00000172915	GO:0005737	cytoplasm	*NBEA*
ENSG00000162599	GO:0006355	regulation of transcription, DNA-dependent	*NFIA*
ENSG00000162599	GO:0006260	DNA replication	*NFIA*
ENSG00000162599	GO:0005634	nucleus	*NFIA*
ENSG00000162599	GO:0005622	intracellular	*NFIA*
ENSG00000162599	GO:0030054	cell junction	*NFIA*
ENSG00000162599	GO:0003677	DNA binding	*NFIA*
ENSG00000162599	GO:0003700	sequence-specific DNA binding transcription factor activity	*NFIA*
ENSG00000162599	GO:0045944	positive regulation of transcription from RNA polymerase II promoter	*NFIA*
ENSG00000162599	GO:0014070	response to organic cyclic compound	*NFIA*
ENSG00000162599	GO:0019079	viral genome replication	*NFIA*
ENSG00000162599	GO:0008134	transcription factor binding	*NFIA*
ENSG00000162599	GO:0045893	positive regulation of transcription, DNA-dependent	*NFIA*
ENSG00000162599	GO:0005515	protein binding	*NFIA*
ENSG00000187258	GO:0007186	G-protein coupled receptor signaling pathway	*NPSR1*
ENSG00000187258	GO:0016021	integral to membrane	*NPSR1*
ENSG00000187258	GO:0005737	cytoplasm	*NPSR1*
ENSG00000187258	GO:0005886	plasma membrane	*NPSR1*
ENSG00000187258	GO:0004872	receptor activity	*NPSR1*
ENSG00000187258	GO:0004871	signal transducer activity	*NPSR1*
ENSG00000187258	GO:0004930	G-protein coupled receptor activity	*NPSR1*
ENSG00000187258	GO:0005000	vasopressin receptor activity	*NPSR1*
ENSG00000122585	GO:0008283	cell proliferation	*NPY*
ENSG00000122585	GO:0007218	neuropeptide signaling pathway	*NPY*
ENSG00000122585	GO:0007586	digestion	*NPY*
ENSG00000122585	GO:0007631	feeding behavior	*NPY*
ENSG00000122585	GO:0006816	calcium ion transport	*NPY*
ENSG00000122585	GO:0021987	cerebral cortex development	*NPY*
ENSG00000122585	GO:0007610	behavior	*NPY*
ENSG00000122585	GO:0007268	synaptic transmission	*NPY*
ENSG00000122585	GO:0045776	negative regulation of blood pressure	*NPY*
ENSG00000122585	GO:0006928	cellular component movement	*NPY*
ENSG00000122585	GO:0050909	sensory perception of taste	*NPY*
ENSG00000122585	GO:0031175	neuron projection development	*NPY*
ENSG00000122585	GO:0021954	central nervous system neuron development	*NPY*
ENSG00000122585	GO:0008343	adult feeding behavior	*NPY*
ENSG00000122585	GO:0032100	positive regulation of appetite	*NPY*
ENSG00000122585	GO:0008015	blood circulation	*NPY*
ENSG00000122585	GO:0007187	G-protein signaling, coupled to cyclic nucleotide second messenger	*NPY*
ENSG00000122585	GO:0008217	regulation of blood pressure	*NPY*
ENSG00000122585	GO:0005576	extracellular region	*NPY*
ENSG00000122585	GO:0005615	extracellular space	*NPY*
ENSG00000122585	GO:0005623	cell	*NPY*
ENSG00000122585	GO:0005102	receptor binding	*NPY*
ENSG00000122585	GO:0004930	G-protein coupled receptor activity	*NPY*
ENSG00000122585	GO:0005179	hormone activity	*NPY*
ENSG00000122585	GO:0001664	G-protein coupled receptor binding	*NPY*
ENSG00000122585	GO:0005246	calcium channel regulator activity	*NPY*

ENSG00000122585	GO:0005184	neuropeptide hormone activity	*NPY*
ENSG00000122585	GO:0031841	neuropeptide Y receptor binding	*NPY*
ENSG00000197893	GO:0008270	zinc ion binding	*NRAP*
ENSG00000197893	GO:0008150	biological_process	*NRAP*
ENSG00000197893	GO:0005916	fascia adherens	*NRAP*
ENSG00000197893	GO:0005927	muscle tendon junction	*NRAP*
ENSG00000197893	GO:0030016	myofibril	*NRAP*
ENSG00000197893	GO:0046872	metal ion binding	*NRAP*
ENSG00000197893	GO:0003779	actin binding	*NRAP*
ENSG00000197893	GO:0051371	muscle alpha-actinin binding	*NRAP*
ENSG00000197893	GO:0017166	vinculin binding	*NRAP*
ENSG00000197893	GO:0005515	protein binding	*NRAP*
ENSG00000021645	GO:0005515	protein binding	*NRXN3*
ENSG00000021645	GO:0007155	cell adhesion	*NRXN3*
ENSG00000021645	GO:0001525	angiogenesis	*NRXN3*
ENSG00000021645	GO:0016020	membrane	*NRXN3*
ENSG00000021645	GO:0016021	integral to membrane	*NRXN3*
ENSG00000021645	GO:0007411	axon guidance	*NRXN3*
ENSG00000021645	GO:0005887	integral to plasma membrane	*NRXN3*
ENSG00000021645	GO:0004872	receptor activity	*NRXN3*
ENSG00000021645	GO:0046872	metal ion binding	*NRXN3*
ENSG00000144460			*NYAP2*
ENSG00000180988	GO:0007186	G-protein coupled receptor signaling pathway	*OR52N2*
ENSG00000180988	GO:0050896	response to stimulus	*OR52N2*
ENSG00000180988	GO:0016021	integral to membrane	*OR52N2*
ENSG00000180988	GO:0005886	plasma membrane	*OR52N2*
ENSG00000180988	GO:0004872	receptor activity	*OR52N2*
ENSG00000180988	GO:0004871	signal transducer activity	*OR52N2*
ENSG00000180988	GO:0004930	G-protein coupled receptor activity	*OR52N2*
ENSG00000180988	GO:0004984	olfactory receptor activity	*OR52N2*
ENSG00000181074	GO:0007186	G-protein coupled receptor signaling pathway	*OR52N4*
ENSG00000181074	GO:0050896	response to stimulus	*OR52N4*
ENSG00000181074	GO:0016021	integral to membrane	*OR52N4*
ENSG00000181074	GO:0005886	plasma membrane	*OR52N4*
ENSG00000181074	GO:0004872	receptor activity	*OR52N4*
ENSG00000181074	GO:0004871	signal transducer activity	*OR52N4*
ENSG00000181074	GO:0004930	G-protein coupled receptor activity	*OR52N4*
ENSG00000181074	GO:0004984	olfactory receptor activity	*OR52N4*
ENSG00000172457	GO:0007186	G-protein coupled receptor signaling pathway	*OR9G4*
ENSG00000172457	GO:0050896	response to stimulus	*OR9G4*
ENSG00000172457	GO:0016021	integral to membrane	*OR9G4*
ENSG00000172457	GO:0005886	plasma membrane	*OR9G4*
ENSG00000172457	GO:0004872	receptor activity	*OR9G4*
ENSG00000172457	GO:0004871	signal transducer activity	*OR9G4*
ENSG00000172457	GO:0004930	G-protein coupled receptor activity	*OR9G4*
ENSG00000172457	GO:0004984	olfactory receptor activity	*OR9G4*
ENSG00000262647	GO:0007186	G-protein coupled receptor signaling pathway	*OR9G4*
ENSG00000262647	GO:0050896	response to stimulus	*OR9G4*
ENSG00000262647	GO:0016021	integral to membrane	*OR9G4*
ENSG00000262647	GO:0005886	plasma membrane	*OR9G4*
ENSG00000262647	GO:0004872	receptor activity	*OR9G4*
ENSG00000262647	GO:0004871	signal transducer activity	*OR9G4*

ENSG00000262647	GO:0004930	G-protein coupled receptor activity	*OR9G4*
ENSG00000262647	GO:0004984	olfactory receptor activity	*OR9G4*
ENSG00000169918	GO:0006508	proteolysis	*OTUD7A*
ENSG00000169918	GO:0005634	nucleus	*OTUD7A*
ENSG00000169918	GO:0005737	cytoplasm	*OTUD7A*
ENSG00000169918	GO:0003677	DNA binding	*OTUD7A*
ENSG00000169918	GO:0008270	zinc ion binding	*OTUD7A*
ENSG00000169918	GO:0046872	metal ion binding	*OTUD7A*
ENSG00000169918	GO:0008233	peptidase activity	*OTUD7A*
ENSG00000169918	GO:0008234	cysteine-type peptidase activity	*OTUD7A*
ENSG00000134853	GO:0007275	multicellular organismal development	*PDGFRA*
ENSG00000134853	GO:0008284	positive regulation of cell proliferation	*PDGFRA*
ENSG00000134853	GO:0042475	odontogenesis of dentin-containing tooth	*PDGFRA*
ENSG00000134853	GO:0042060	wound healing	*PDGFRA*
ENSG00000134853	GO:0048839	inner ear development	*PDGFRA*
ENSG00000134853	GO:0018108	peptidyl-tyrosine phosphorylation	*PDGFRA*
ENSG00000134853	GO:0006468	protein phosphorylation	*PDGFRA*
ENSG00000134853	GO:0007169	transmembrane receptor protein tyrosine kinase signaling pathway	*PDGFRA*
ENSG00000134853	GO:0032355	response to estradiol stimulus	*PDGFRA*
ENSG00000134853	GO:0007204	elevation of cytosolic calcium ion concentration	*PDGFRA*
ENSG00000134853	GO:0048015	phosphatidylinositol-mediated signaling	*PDGFRA*
ENSG00000134853	GO:0030324	lung development	*PDGFRA*
ENSG00000134853	GO:0048146	positive regulation of fibroblast proliferation	*PDGFRA*
ENSG00000134853	GO:0008585	female gonad development	*PDGFRA*
ENSG00000134853	GO:0001701	in utero embryonic development	*PDGFRA*
ENSG00000134853	GO:0048705	skeletal system morphogenesis	*PDGFRA*
ENSG00000134853	GO:0055003	cardiac myofibril assembly	*PDGFRA*
ENSG00000134853	GO:0030198	extracellular matrix organization	*PDGFRA*
ENSG00000134853	GO:0034097	response to cytokine stimulus	*PDGFRA*
ENSG00000134853	GO:0070374	positive regulation of ERK1 and ERK2 cascade	*PDGFRA*
ENSG00000134853	GO:0048704	embryonic skeletal system morphogenesis	*PDGFRA*
ENSG00000134853	GO:0030335	positive regulation of cell migration	*PDGFRA*
ENSG00000134853	GO:0048008	platelet-derived growth factor receptor signaling pathway	*PDGFRA*
ENSG00000134853	GO:0009725	response to hormone stimulus	*PDGFRA*
ENSG00000134853	GO:0001775	cell activation	*PDGFRA*
ENSG00000134853	GO:0071230	cellular response to amino acid stimulus	*PDGFRA*
ENSG00000134853	GO:0030325	adrenal gland development	*PDGFRA*
ENSG00000134853	GO:0046777	protein autophosphorylation	*PDGFRA*
ENSG00000134853	GO:0055093	response to hyperoxia	*PDGFRA*
ENSG00000134853	GO:0023019	signal transduction involved in regulation of gene expression	*PDGFRA*
ENSG00000134853	GO:0045740	positive regulation of DNA replication	*PDGFRA*
ENSG00000134853	GO:0010035	response to inorganic substance	*PDGFRA*
ENSG00000134853	GO:0060326	cell chemotaxis	*PDGFRA*
ENSG00000134853	GO:0048557	embryonic digestive tract morphogenesis	*PDGFRA*
ENSG00000134853	GO:0060021	palate development	*PDGFRA*
ENSG00000134853	GO:0030539	male genitalia development	*PDGFRA*
ENSG00000134853	GO:0010544	negative regulation of platelet activation	*PDGFRA*
ENSG00000134853	GO:0014068	positive regulation of phosphatidylinositol 3-kinase cascade	*PDGFRA*
ENSG00000134853	GO:0060325	face morphogenesis	*PDGFRA*
ENSG00000134853	GO:0048701	embryonic cranial skeleton morphogenesis	*PDGFRA*

ENSG00000134853	GO:0061298	retina vasculature development in camera-type eye	*PDGFRA*
ENSG00000134853	GO:0033327	Leydig cell differentiation	*PDGFRA*
ENSG00000134853	GO:0001553	luteinization	*PDGFRA*
ENSG00000134853	GO:0008210	estrogen metabolic process	*PDGFRA*
ENSG00000134853	GO:0042063	gliogenesis	*PDGFRA*
ENSG00000134853	GO:0072277	metanephric glomerular capillary formation	*PDGFRA*
ENSG00000134853	GO:0048010	vascular endothelial growth factor receptor signaling pathway	*PDGFRA*
ENSG00000134853	GO:2000249	regulation of actin cytoskeleton reorganization	*PDGFRA*
ENSG00000134853	GO:0043552	positive regulation of phosphatidylinositol 3-kinase activity	*PDGFRA*
ENSG00000134853	GO:0070527	platelet aggregation	*PDGFRA*
ENSG00000134853	GO:0010863	positive regulation of phospholipase C activity	*PDGFRA*
ENSG00000134853	GO:0038091	positive regulation of cell proliferation by VEGF-activated platelet derived growth factor receptor signaling pathway	*PDGFRA*
ENSG00000134853	GO:0035790	platelet-derived growth factor receptor-alpha signaling pathway	*PDGFRA*
ENSG00000134853	GO:0050920	regulation of chemotaxis	*PDGFRA*
ENSG00000134853	GO:2000739	regulation of mesenchymal stem cell differentiation	*PDGFRA*
ENSG00000134853	GO:0010033	response to organic substance	*PDGFRA*
ENSG00000134853	GO:0043627	response to estrogen stimulus	*PDGFRA*
ENSG00000134853	GO:0009653	anatomical structure morphogenesis	*PDGFRA*
ENSG00000134853	GO:0009887	organ morphogenesis	*PDGFRA*
ENSG00000134853	GO:0016477	cell migration	*PDGFRA*
ENSG00000134853	GO:0016021	integral to membrane	*PDGFRA*
ENSG00000134853	GO:0005634	nucleus	*PDGFRA*
ENSG00000134853	GO:0005737	cytoplasm	*PDGFRA*
ENSG00000134853	GO:0005886	plasma membrane	*PDGFRA*
ENSG00000134853	GO:0005887	integral to plasma membrane	*PDGFRA*
ENSG00000134853	GO:0031226	intrinsic to plasma membrane	*PDGFRA*
ENSG00000134853	GO:0004872	receptor activity	*PDGFRA*
ENSG00000134853	GO:0000166	nucleotide binding	*PDGFRA*
ENSG00000134853	GO:0042803	protein homodimerization activity	*PDGFRA*
ENSG00000134853	GO:0005515	protein binding	*PDGFRA*
ENSG00000134853	GO:0005524	ATP binding	*PDGFRA*
ENSG00000134853	GO:0004672	protein kinase activity	*PDGFRA*
ENSG00000134853	GO:0004714	transmembrane receptor protein tyrosine kinase activity	*PDGFRA*
ENSG00000134853	GO:0004713	protein tyrosine kinase activity	*PDGFRA*
ENSG00000134853	GO:0005161	platelet-derived growth factor receptor binding	*PDGFRA*
ENSG00000134853	GO:0043548	phosphatidylinositol 3-kinase binding	*PDGFRA*
ENSG00000134853	GO:0005018	platelet-derived growth factor alpha-receptor activity	*PDGFRA*
ENSG00000134853	GO:0048407	platelet-derived growth factor binding	*PDGFRA*
ENSG00000134853	GO:0005021	vascular endothelial growth factor-activated receptor activity	*PDGFRA*
ENSG00000134853	GO:0038085	vascular endothelial growth factor binding	*PDGFRA*
ENSG00000182405			*PGBD4*
ENSG00000079739	GO:0044281	small molecule metabolic process	*PGM1*
ENSG00000079739	GO:0005975	carbohydrate metabolic process	*PGM1*
ENSG00000079739	GO:0006006	glucose metabolic process	*PGM1*

ENSG00000079739	GO:0005978	glycogen biosynthetic process	*PGM1*
ENSG00000079739	GO:0019255	glucose 1-phosphate metabolic process	*PGM1*
ENSG00000079739	GO:0019388	galactose catabolic process	*PGM1*
ENSG00000079739	GO:0005980	glycogen catabolic process	*PGM1*
ENSG00000079739	GO:0005737	cytoplasm	*PGM1*
ENSG00000079739	GO:0005829	cytosol	*PGM1*
ENSG00000079739	GO:0015629	actin cytoskeleton	*PGM1*
ENSG00000079739	GO:0000287	magnesium ion binding	*PGM1*
ENSG00000079739	GO:0016853	isomerase activity	*PGM1*
ENSG00000079739	GO:0004614	phosphoglucomutase activity	*PGM1*
ENSG00000079739	GO:0016868	intramolecular transferase activity, phosphotransferases	*PGM1*
ENSG00000122861	GO:0006508	proteolysis	*PLAU*
ENSG00000122861	GO:0005515	protein binding	*PLAU*
ENSG00000122861	GO:0004252	serine-type endopeptidase activity	*PLAU*
ENSG00000122861	GO:0007165	signal transduction	*PLAU*
ENSG00000122861	GO:0007566	embryo implantation	*PLAU*
ENSG00000122861	GO:0001525	angiogenesis	*PLAU*
ENSG00000122861	GO:0042730	fibrinolysis	*PLAU*
ENSG00000122861	GO:0001666	response to hypoxia	*PLAU*
ENSG00000122861	GO:0042127	regulation of cell proliferation	*PLAU*
ENSG00000122861	GO:0006935	chemotaxis	*PLAU*
ENSG00000122861	GO:0055093	response to hyperoxia	*PLAU*
ENSG00000122861	GO:0007596	blood coagulation	*PLAU*
ENSG00000122861	GO:0033628	regulation of cell adhesion mediated by integrin	*PLAU*
ENSG00000122861	GO:0061041	regulation of wound healing	*PLAU*
ENSG00000122861	GO:0043403	skeletal muscle tissue regeneration	*PLAU*
ENSG00000122861	GO:0010469	regulation of receptor activity	*PLAU*
ENSG00000122861	GO:0014910	regulation of smooth muscle cell migration	*PLAU*
ENSG00000122861	GO:2000097	regulation of smooth muscle cell-matrix adhesion	*PLAU*
ENSG00000122861	GO:0042060	wound healing	*PLAU*
ENSG00000122861	GO:0048514	blood vessel morphogenesis	*PLAU*
ENSG00000122861	GO:0014909	smooth muscle cell migration	*PLAU*
ENSG00000122861	GO:0005886	plasma membrane	*PLAU*
ENSG00000122861	GO:0005576	extracellular region	*PLAU*
ENSG00000122861	GO:0005615	extracellular space	*PLAU*
ENSG00000122861	GO:0009986	cell surface	*PLAU*
ENSG00000122861	GO:0008233	peptidase activity	*PLAU*
ENSG00000122861	GO:0016301	kinase activity	*PLAU*
ENSG00000100941	GO:0007155	cell adhesion	*PNN*
ENSG00000100941	GO:0006355	regulation of transcription, DNA-dependent	*PNN*
ENSG00000100941	GO:0008380	RNA splicing	*PNN*
ENSG00000100941	GO:0000398	nuclear mRNA splicing, via spliceosome	*PNN*
ENSG00000100941	GO:0005634	nucleus	*PNN*
ENSG00000100941	GO:0005737	cytoplasm	*PNN*
ENSG00000100941	GO:0005886	plasma membrane	*PNN*
ENSG00000100941	GO:0005882	intermediate filament	*PNN*
ENSG00000100941	GO:0005911	cell-cell junction	*PNN*
ENSG00000100941	GO:0016607	nuclear speck	*PNN*
ENSG00000100941	GO:0030057	desmosome	*PNN*
ENSG00000100941	GO:0071013	catalytic step 2 spliceosome	*PNN*
ENSG00000100941	GO:0003677	DNA binding	*PNN*
ENSG00000100941	GO:0005198	structural molecule activity	*PNN*
ENSG00000128567	GO:0030335	positive regulation of cell migration	*PODXL*

ENSG00000128567	GO:0016477	cell migration	*PODXL*
ENSG00000128567	GO:0007162	negative regulation of cell adhesion	*PODXL*
ENSG00000128567	GO:0050900	leukocyte migration	*PODXL*
ENSG00000128567	GO:0022408	negative regulation of cell-cell adhesion	*PODXL*
ENSG00000128567	GO:0032534	regulation of microvillus assembly	*PODXL*
ENSG00000128567	GO:0033634	positive regulation of cell-cell adhesion mediated by integrin	*PODXL*
ENSG00000128567	GO:0072015	glomerular visceral epithelial cell development	*PODXL*
ENSG00000128567	GO:0072175	epithelial tube formation	*PODXL*
ENSG00000128567	GO:0005737	cytoplasm	*PODXL*
ENSG00000128567	GO:0005886	plasma membrane	*PODXL*
ENSG00000128567	GO:0005730	nucleolus	*PODXL*
ENSG00000128567	GO:0043231	intracellular membrane-bounded organelle	*PODXL*
ENSG00000128567	GO:0005813	centrosome	*PODXL*
ENSG00000128567	GO:0015629	actin cytoskeleton	*PODXL*
ENSG00000128567	GO:0001726	ruffle	*PODXL*
ENSG00000128567	GO:0005887	integral to plasma membrane	*PODXL*
ENSG00000128567	GO:0030027	lamellipodium	*PODXL*
ENSG00000128567	GO:0016324	apical plasma membrane	*PODXL*
ENSG00000128567	GO:0045121	membrane raft	*PODXL*
ENSG00000128567	GO:0030175	filopodium	*PODXL*
ENSG00000128567	GO:0031528	microvillus membrane	*PODXL*
ENSG00000128567	GO:0036057	slit diaphragm	*PODXL*
ENSG00000128567	GO:0005515	protein binding	*PODXL*
ENSG00000138738	GO:0045892	negative regulation of transcription, DNA-dependent	*PRDM5*
ENSG00000138738	GO:0000278	mitotic cell cycle	*PRDM5*
ENSG00000138738	GO:0016568	chromatin modification	*PRDM5*
ENSG00000138738	GO:0051567	histone H3-K9 methylation	*PRDM5*
ENSG00000138738	GO:0016575	histone deacetylation	*PRDM5*
ENSG00000138738	GO:0005634	nucleus	*PRDM5*
ENSG00000138738	GO:0005622	intracellular	*PRDM5*
ENSG00000138738	GO:0008270	zinc ion binding	*PRDM5*
ENSG00000138738	GO:0005515	protein binding	*PRDM5*
ENSG00000138738	GO:0046872	metal ion binding	*PRDM5*
ENSG00000138738	GO:0043565	sequence-specific DNA binding	*PRDM5*
ENSG00000138738	GO:0000976	transcription regulatory region sequence-specific DNA binding	*PRDM5*
ENSG00000138738	GO:0044212	transcription regulatory region DNA binding	*PRDM5*
ENSG00000138738	GO:0070491	repressing transcription factor binding	*PRDM5*
ENSG00000138738	GO:0003676	nucleic acid binding	*PRDM5*
ENSG00000164244	GO:0005794	Golgi apparatus	*PRRC1*
ENSG00000078328	GO:0000166	nucleotide binding	*RBFOX1*
ENSG00000078328	GO:0003676	nucleic acid binding	*RBFOX1*
ENSG00000078328	GO:0006397	mRNA processing	*RBFOX1*
ENSG00000078328	GO:0008380	RNA splicing	*RBFOX1*
ENSG00000078328	GO:0043484	regulation of RNA splicing	*RBFOX1*
ENSG00000078328	GO:0050658	RNA transport	*RBFOX1*
ENSG00000078328	GO:0005634	nucleus	*RBFOX1*
ENSG00000078328	GO:0005737	cytoplasm	*RBFOX1*
ENSG00000078328	GO:0005802	trans-Golgi network	*RBFOX1*
ENSG00000078328	GO:0003723	RNA binding	*RBFOX1*
ENSG00000078328	GO:0005515	protein binding	*RBFOX1*
ENSG00000078328	GO:0008022	protein C-terminus binding	*RBFOX1*
ENSG00000143839	GO:0006508	proteolysis	*REN*

ENSG00000143839	GO:0006950	response to stress	*REN*
ENSG00000143839	GO:0008584	male gonad development	*REN*
ENSG00000143839	GO:0042493	response to drug	*REN*
ENSG00000143839	GO:0048469	cell maturation	*REN*
ENSG00000143839	GO:0051591	response to cAMP	*REN*
ENSG00000143839	GO:0008217	regulation of blood pressure	*REN*
ENSG00000143839	GO:0001822	kidney development	*REN*
ENSG00000143839	GO:0043408	regulation of MAPK cascade	*REN*
ENSG00000143839	GO:0001823	mesonephros development	*REN*
ENSG00000143839	GO:0009755	hormone-mediated signaling pathway	*REN*
ENSG00000143839	GO:0002003	angiotensin maturation	*REN*
ENSG00000143839	GO:0002018	renin-angiotensin regulation of aldosterone production	*REN*
ENSG00000143839	GO:0042756	drinking behavior	*REN*
ENSG00000143839	GO:0070305	response to cGMP	*REN*
ENSG00000143839	GO:0002016	regulation of blood volume by renin-angiotensin	*REN*
ENSG00000143839	GO:0010033	response to organic substance	*REN*
ENSG00000143839	GO:0016020	membrane	*REN*
ENSG00000143839	GO:0005576	extracellular region	*REN*
ENSG00000143839	GO:0005622	intracellular	*REN*
ENSG00000143839	GO:0005615	extracellular space	*REN*
ENSG00000143839	GO:0044444	cytoplasmic part	*REN*
ENSG00000143839	GO:0005102	receptor binding	*REN*
ENSG00000143839	GO:0008233	peptidase activity	*REN*
ENSG00000143839	GO:0004190	aspartic-type endopeptidase activity	*REN*
ENSG00000143839	GO:0005159	insulin-like growth factor receptor binding	*REN*
ENSG00000143839	GO:0004175	endopeptidase activity	*REN*
ENSG00000152214	GO:0007165	signal transduction	*RIT2*
ENSG00000152214	GO:0007264	small GTPase mediated signal transduction	*RIT2*
ENSG00000152214	GO:0006886	intracellular protein transport	*RIT2*
ENSG00000152214	GO:0015031	protein transport	*RIT2*
ENSG00000152214	GO:0006913	nucleocytoplasmic transport	*RIT2*
ENSG00000152214	GO:0007268	synaptic transmission	*RIT2*
ENSG00000152214	GO:0006184	GTP catabolic process	*RIT2*
ENSG00000152214	GO:0048011	nerve growth factor receptor signaling pathway	*RIT2*
ENSG00000152214	GO:0016020	membrane	*RIT2*
ENSG00000152214	GO:0005886	plasma membrane	*RIT2*
ENSG00000152214	GO:0005622	intracellular	*RIT2*
ENSG00000152214	GO:0000166	nucleotide binding	*RIT2*
ENSG00000152214	GO:0005516	calmodulin binding	*RIT2*
ENSG00000152214	GO:0005525	GTP binding	*RIT2*
ENSG00000152214	GO:0003924	GTPase activity	*RIT2*
ENSG00000152214	GO:0005515	protein binding	*RIT2*
ENSG00000185008	GO:0007411	axon guidance	*ROBO2*
ENSG00000185008	GO:0001656	metanephros development	*ROBO2*
ENSG00000185008	GO:0030673	axolemma	*ROBO2*
ENSG00000185008	GO:0005515	protein binding	*ROBO2*
ENSG00000185008	GO:0046982	protein heterodimerization activity	*ROBO2*
ENSG00000185008	GO:0007275	multicellular organismal development	*ROBO2*
ENSG00000185008	GO:0030154	cell differentiation	*ROBO2*
ENSG00000185008	GO:0007420	brain development	*ROBO2*
ENSG00000185008	GO:0007417	central nervous system development	*ROBO2*
ENSG00000185008	GO:0007156	homophilic cell adhesion	*ROBO2*
ENSG00000185008	GO:0050772	positive regulation of axonogenesis	*ROBO2*

ENSG00000185008	GO:0032870	cellular response to hormone stimulus	*ROBO2*
ENSG00000185008	GO:0001657	ureteric bud development	*ROBO2*
ENSG00000185008	GO:0031290	retinal ganglion cell axon guidance	*ROBO2*
ENSG00000185008	GO:0051964	negative regulation of synapse assembly	*ROBO2*
ENSG00000185008	GO:0021891	olfactory bulb interneuron development	*ROBO2*
ENSG00000185008	GO:0016199	axon midline choice point recognition	*ROBO2*
ENSG00000185008	GO:0061364	apoptotic process involved in luteolysis	*ROBO2*
ENSG00000185008	GO:0050925	negative regulation of negative chemotaxis	*ROBO2*
ENSG00000185008	GO:0016020	membrane	*ROBO2*
ENSG00000185008	GO:0016021	integral to membrane	*ROBO2*
ENSG00000185008	GO:0009986	cell surface	*ROBO2*
ENSG00000185008	GO:0042802	identical protein binding	*ROBO2*
ENSG00000185008	GO:0008046	axon guidance receptor activity	*ROBO2*
ENSG00000185008	GO:0007155	cell adhesion	*ROBO2*
ENSG00000143365	GO:0006355	regulation of transcription, DNA-dependent	*RORC*
ENSG00000143365	GO:0043401	steroid hormone mediated signaling pathway	*RORC*
ENSG00000143365	GO:0005634	nucleus	*RORC*
ENSG00000143365	GO:0003700	sequence-specific DNA binding transcription factor activity	*RORC*
ENSG00000143365	GO:0008270	zinc ion binding	*RORC*
ENSG00000143365	GO:0043565	sequence-specific DNA binding	*RORC*
ENSG00000143365	GO:0003707	steroid hormone receptor activity	*RORC*
ENSG00000143365	GO:0007275	multicellular organismal development	*RORC*
ENSG00000143365	GO:0030154	cell differentiation	*RORC*
ENSG00000143365	GO:0006468	protein phosphorylation	*RORC*
ENSG00000143365	GO:0010467	gene expression	*RORC*
ENSG00000143365	GO:0033077	T cell differentiation in thymus	*RORC*
ENSG00000143365	GO:0046632	alpha-beta T cell differentiation	*RORC*
ENSG00000143365	GO:0048535	lymph node development	*RORC*
ENSG00000143365	GO:0030522	intracellular receptor mediated signaling pathway	*RORC*
ENSG00000143365	GO:0048537	mucosal-associated lymphoid tissue development	*RORC*
ENSG00000143365	GO:0072539	T-helper 17 cell differentiation	*RORC*
ENSG00000143365	GO:0005654	nucleoplasm	*RORC*
ENSG00000143365	GO:0004872	receptor activity	*RORC*
ENSG00000143365	GO:0003677	DNA binding	*RORC*
ENSG00000143365	GO:0046872	metal ion binding	*RORC*
ENSG00000143365	GO:0004879	ligand-activated sequence-specific DNA binding RNA polymerase II transcription factor activity	*RORC*
ENSG00000143365	GO:0005515	protein binding	*RORC*
ENSG00000114993	GO:0007165	signal transduction	*RTKN*
ENSG00000114993	GO:0005622	intracellular	*RTKN*
ENSG00000114993	GO:0005515	protein binding	*RTKN*
ENSG00000114993	GO:0005543	phospholipid binding	*RTKN*
ENSG00000114993	GO:0006915	apoptotic process	*RTKN*
ENSG00000114993	GO:0007266	Rho protein signal transduction	*RTKN*
ENSG00000114993	GO:0045767	regulation of anti-apoptosis	*RTKN*
ENSG00000114993	GO:0005575	cellular_component	*RTKN*
ENSG00000114993	GO:0000166	nucleotide binding	*RTKN*
ENSG00000114993	GO:0005525	GTP binding	*RTKN*
ENSG00000114993	GO:0005095	GTPase inhibitor activity	*RTKN*
ENSG00000114993	GO:0017049	GTP-Rho binding	*RTKN*
ENSG00000114993	GO:0017048	Rho GTPase binding	*RTKN*
ENSG00000075223	GO:0007411	axon guidance	*SEMA3C*

ENSG00000075223	GO:0030154	cell differentiation	*SEMA3C*
ENSG00000075223	GO:0007507	heart development	*SEMA3C*
ENSG00000075223	GO:0001974	blood vessel remodeling	*SEMA3C*
ENSG00000075223	GO:0006955	immune response	*SEMA3C*
ENSG00000075223	GO:0009791	post-embryonic development	*SEMA3C*
ENSG00000075223	GO:0042493	response to drug	*SEMA3C*
ENSG00000075223	GO:0001756	somitogenesis	*SEMA3C*
ENSG00000075223	GO:0001755	neural crest cell migration	*SEMA3C*
ENSG00000075223	GO:0003151	outflow tract morphogenesis	*SEMA3C*
ENSG00000075223	GO:0021915	neural tube development	*SEMA3C*
ENSG00000075223	GO:0060666	dichotomous subdivision of terminal units involved in salivary gland branching	*SEMA3C*
ENSG00000075223	GO:0003215	cardiac right ventricle morphogenesis	*SEMA3C*
ENSG00000075223	GO:0060174	limb bud formation	*SEMA3C*
ENSG00000075223	GO:0003350	pulmonary myocardium development	*SEMA3C*
ENSG00000075223	GO:0016020	membrane	*SEMA3C*
ENSG00000075223	GO:0005576	extracellular region	*SEMA3C*
ENSG00000075223	GO:0005615	extracellular space	*SEMA3C*
ENSG00000075223	GO:0004872	receptor activity	*SEMA3C*
ENSG00000075223	GO:0005515	protein binding	*SEMA3C*
ENSG00000075223	GO:0030215	semaphorin receptor binding	*SEMA3C*
ENSG00000075223	GO:0007275	multicellular organismal development	*SEMA3C*
ENSG00000155926	GO:0005737	cytoplasm	*SLA*
ENSG00000155926	GO:0005768	endosome	*SLA*
ENSG00000155926	GO:0005515	protein binding	*SLA*
ENSG00000155926	GO:0005070	SH3/SH2 adaptor activity	*SLA*
ENSG00000109171	GO:0031116	positive regulation of microtubule polymerization	*SLAIN2*
ENSG00000109171	GO:0031122	cytoplasmic microtubule organization	*SLAIN2*
ENSG00000109171	GO:0007020	microtubule nucleation	*SLAIN2*
ENSG00000109171	GO:0031113	regulation of microtubule polymerization	*SLAIN2*
ENSG00000109171	GO:0005737	cytoplasm	*SLAIN2*
ENSG00000109171	GO:0005813	centrosome	*SLAIN2*
ENSG00000109171	GO:0015630	microtubule cytoskeleton	*SLAIN2*
ENSG00000109171	GO:0035371	microtubule plus end	*SLAIN2*
ENSG00000109171	GO:0005515	protein binding	*SLAIN2*
ENSG00000124140	GO:0006811	ion transport	*SLC12A5*
ENSG00000124140	GO:0006813	potassium ion transport	*SLC12A5*
ENSG00000124140	GO:0055085	transmembrane transport	*SLC12A5*
ENSG00000124140	GO:0035264	multicellular organism growth	*SLC12A5*
ENSG00000124140	GO:0006810	transport	*SLC12A5*
ENSG00000124140	GO:0042493	response to drug	*SLC12A5*
ENSG00000124140	GO:0007268	synaptic transmission	*SLC12A5*
ENSG00000124140	GO:0006821	chloride transport	*SLC12A5*
ENSG00000124140	GO:0007612	learning	*SLC12A5*
ENSG00000124140	GO:0051452	intracellular pH reduction	*SLC12A5*
ENSG00000124140	GO:0015696	ammonium transport	*SLC12A5*
ENSG00000124140	GO:0040040	thermosensory behavior	*SLC12A5*
ENSG00000124140	GO:0016020	membrane	*SLC12A5*
ENSG00000124140	GO:0016021	integral to membrane	*SLC12A5*
ENSG00000124140	GO:0005886	plasma membrane	*SLC12A5*
ENSG00000124140	GO:0043204	perikaryon	*SLC12A5*
ENSG00000124140	GO:0032590	dendrite membrane	*SLC12A5*
ENSG00000124140	GO:0015379	potassium:chloride symporter activity	*SLC12A5*
ENSG00000124140	GO:0051739	ammonia transmembrane transporter activity	*SLC12A5*

ENSG00000124140	GO:0015377	cation:chloride symporter activity	*SLC12A5*
ENSG00000124140	GO:0006873	cellular ion homeostasis	*SLC12A5*
ENSG00000124140	GO:0015293	symporter activity	*SLC12A5*
ENSG00000124140	GO:0005515	protein binding	*SLC12A5*
ENSG00000109762	GO:0007154	cell communication	*SNX25*
ENSG00000109762	GO:0015031	protein transport	*SNX25*
ENSG00000109762	GO:0038032	termination of G-protein coupled receptor signaling pathway	*SNX25*
ENSG00000109762	GO:0016020	membrane	*SNX25*
ENSG00000109762	GO:0005768	endosome	*SNX25*
ENSG00000109762	GO:0010008	endosome membrane	*SNX25*
ENSG00000109762	GO:0005515	protein binding	*SNX25*
ENSG00000109762	GO:0035091	phosphatidylinositol binding	*SNX25*
ENSG00000130340	GO:0006886	intracellular protein transport	*SNX9*
ENSG00000130340	GO:0007154	cell communication	*SNX9*
ENSG00000130340	GO:0006898	receptor-mediated endocytosis	*SNX9*
ENSG00000130340	GO:0043547	positive regulation of GTPase activity	*SNX9*
ENSG00000130340	GO:0032461	positive regulation of protein oligomerization	*SNX9*
ENSG00000130340	GO:0060988	lipid tube assembly	*SNX9*
ENSG00000130340	GO:0005737	cytoplasm	*SNX9*
ENSG00000130340	GO:0005886	plasma membrane	*SNX9*
ENSG00000130340	GO:0005622	intracellular	*SNX9*
ENSG00000130340	GO:0005794	Golgi apparatus	*SNX9*
ENSG00000130340	GO:0005802	trans-Golgi network	*SNX9*
ENSG00000130340	GO:0001726	ruffle	*SNX9*
ENSG00000130340	GO:0030659	cytoplasmic vesicle membrane	*SNX9*
ENSG00000130340	GO:0005625	soluble fraction	*SNX9*
ENSG00000130340	GO:0005792	microsome	*SNX9*
ENSG00000130340	GO:0031234	extrinsic to internal side of plasma membrane	*SNX9*
ENSG00000130340	GO:0016023	cytoplasmic membrane-bounded vesicle	*SNX9*
ENSG00000130340	GO:0030136	clathrin-coated vesicle	*SNX9*
ENSG00000130340	GO:0042803	protein homodimerization activity	*SNX9*
ENSG00000130340	GO:0005515	protein binding	*SNX9*
ENSG00000130340	GO:0008289	lipid binding	*SNX9*
ENSG00000130340	GO:0031625	ubiquitin protein ligase binding	*SNX9*
ENSG00000130340	GO:0035091	phosphatidylinositol binding	*SNX9*
ENSG00000130340	GO:0005545	1-phosphatidylinositol binding	*SNX9*
ENSG00000124766	GO:0008285	negative regulation of cell proliferation	*SOX4*
ENSG00000124766	GO:0045893	positive regulation of transcription, DNA-dependent	*SOX4*
ENSG00000124766	GO:0008284	positive regulation of cell proliferation	*SOX4*
ENSG00000124766	GO:0043066	negative regulation of apoptotic process	*SOX4*
ENSG00000124766	GO:0045944	positive regulation of transcription from RNA polymerase II promoter	*SOX4*
ENSG00000124766	GO:0001501	skeletal system development	*SOX4*
ENSG00000124766	GO:0007507	heart development	*SOX4*
ENSG00000124766	GO:0006355	regulation of transcription, DNA-dependent	*SOX4*
ENSG00000124766	GO:0048485	sympathetic nervous system development	*SOX4*
ENSG00000124766	GO:0060070	canonical Wnt receptor signaling pathway	*SOX4*
ENSG00000124766	GO:0006977	DNA damage response, signal transduction by p53 class mediator resulting in cell cycle arrest	*SOX4*
ENSG00000124766	GO:0021522	spinal cord motor neuron differentiation	*SOX4*
ENSG00000124766	GO:0043065	positive regulation of apoptotic process	*SOX4*
ENSG00000124766	GO:0042593	glucose homeostasis	*SOX4*

ENSG00000124766	GO:0035019	somatic stem cell maintenance	*SOX4*
ENSG00000124766	GO:0031647	regulation of protein stability	*SOX4*
ENSG00000124766	GO:0032024	positive regulation of insulin secretion	*SOX4*
ENSG00000124766	GO:0090263	positive regulation of canonical Wnt receptor signaling pathway	*SOX4*
ENSG00000124766	GO:0045727	positive regulation of translation	*SOX4*
ENSG00000124766	GO:0060548	negative regulation of cell death	*SOX4*
ENSG00000124766	GO:0060412	ventricular septum morphogenesis	*SOX4*
ENSG00000124766	GO:0001841	neural tube formation	*SOX4*
ENSG00000124766	GO:0050821	protein stabilization	*SOX4*
ENSG00000124766	GO:0031018	endocrine pancreas development	*SOX4*
ENSG00000124766	GO:0071333	cellular response to glucose stimulus	*SOX4*
ENSG00000124766	GO:0021510	spinal cord development	*SOX4*
ENSG00000124766	GO:0030217	T cell differentiation	*SOX4*
ENSG00000124766	GO:0014009	glial cell proliferation	*SOX4*
ENSG00000124766	GO:0030177	positive regulation of Wnt receptor signaling pathway	*SOX4*
ENSG00000124766	GO:0031397	negative regulation of protein ubiquitination	*SOX4*
ENSG00000124766	GO:0003289	atrial septum primum morphogenesis	*SOX4*
ENSG00000124766	GO:0003215	cardiac right ventricle morphogenesis	*SOX4*
ENSG00000124766	GO:0003211	cardiac ventricle formation	*SOX4*
ENSG00000124766	GO:0060174	limb bud formation	*SOX4*
ENSG00000124766	GO:0002328	pro-B cell differentiation	*SOX4*
ENSG00000124766	GO:0003357	noradrenergic neuron differentiation	*SOX4*
ENSG00000124766	GO:0042769	DNA damage response, detection of DNA damage	*SOX4*
ENSG00000124766	GO:0046826	negative regulation of protein export from nucleus	*SOX4*
ENSG00000124766	GO:0003183	mitral valve morphogenesis	*SOX4*
ENSG00000124766	GO:0021782	glial cell development	*SOX4*
ENSG00000124766	GO:0060563	neuroepithelial cell differentiation	*SOX4*
ENSG00000124766	GO:0060993	kidney morphogenesis	*SOX4*
ENSG00000124766	GO:0035910	ascending aorta morphogenesis	*SOX4*
ENSG00000124766	GO:2000761	positive regulation of N-terminal peptidyl-lysine acetylation	*SOX4*
ENSG00000124766	GO:0005634	nucleus	*SOX4*
ENSG00000124766	GO:0005737	cytoplasm	*SOX4*
ENSG00000124766	GO:0005739	mitochondrion	*SOX4*
ENSG00000124766	GO:0003677	DNA binding	*SOX4*
ENSG00000124766	GO:0003700	sequence-specific DNA binding transcription factor activity	*SOX4*
ENSG00000124766	GO:0005515	protein binding	*SOX4*
ENSG00000124766	GO:0000976	transcription regulatory region sequence-specific DNA binding	*SOX4*
ENSG00000124766	GO:0001077	RNA polymerase II core promoter proximal region sequence-specific DNA binding transcription factor activity involved in positive regulation of transcription	*SOX4*
ENSG00000124766	GO:0001105	RNA polymerase II transcription coactivator activity	*SOX4*
ENSG00000124766	GO:0001046	core promoter sequence-specific DNA binding	*SOX4*
ENSG00000124766	GO:0001071	nucleic acid binding transcription factor activity	*SOX4*
ENSG00000144681	GO:0007165	signal transduction	*STAC*
ENSG00000144681	GO:0035556	intracellular signal transduction	*STAC*
ENSG00000144681	GO:0034605	cellular response to heat	*STAC*

ENSG00000144681	GO:0005737	cytoplasm	*STAC*
ENSG00000144681	GO:0005625	soluble fraction	*STAC*
ENSG00000144681	GO:0005515	protein binding	*STAC*
ENSG00000144681	GO:0046872	metal ion binding	*STAC*
ENSG00000198648	GO:0006468	protein phosphorylation	*STK39*
ENSG00000198648	GO:0006950	response to stress	*STK39*
ENSG00000198648	GO:0023014	signal transduction by phosphorylation	*STK39*
ENSG00000198648	GO:0043268	positive regulation of potassium ion transport	*STK39*
ENSG00000198648	GO:0005634	nucleus	*STK39*
ENSG00000198648	GO:0005737	cytoplasm	*STK39*
ENSG00000198648	GO:0005624	membrane fraction	*STK39*
ENSG00000198648	GO:0005856	cytoskeleton	*STK39*
ENSG00000198648	GO:0016323	basolateral plasma membrane	*STK39*
ENSG00000198648	GO:0016324	apical plasma membrane	*STK39*
ENSG00000198648	GO:0000166	nucleotide binding	*STK39*
ENSG00000198648	GO:0005524	ATP binding	*STK39*
ENSG00000198648	GO:0004672	protein kinase activity	*STK39*
ENSG00000198648	GO:0004713	protein tyrosine kinase activity	*STK39*
ENSG00000198648	GO:0019901	protein kinase binding	*STK39*
ENSG00000198648	GO:0004702	receptor signaling protein serine/threonine kinase activity	*STK39*
ENSG00000198648	GO:0004674	protein serine/threonine kinase activity	*STK39*
ENSG00000198648	GO:0005515	protein binding	*STK39*
ENSG00000104435	GO:0030182	neuron differentiation	*STMN2*
ENSG00000104435	GO:0031115	negative regulation of microtubule polymerization	*STMN2*
ENSG00000104435	GO:0035556	intracellular signal transduction	*STMN2*
ENSG00000104435	GO:0007026	negative regulation of microtubule depolymerization	*STMN2*
ENSG00000104435	GO:0010977	negative regulation of neuron projection development	*STMN2*
ENSG00000104435	GO:0010976	positive regulation of neuron projection development	*STMN2*
ENSG00000104435	GO:0031117	positive regulation of microtubule depolymerization	*STMN2*
ENSG00000104435	GO:0016020	membrane	*STMN2*
ENSG00000104435	GO:0005737	cytoplasm	*STMN2*
ENSG00000104435	GO:0030424	axon	*STMN2*
ENSG00000104435	GO:0042995	cell projection	*STMN2*
ENSG00000104435	GO:0005624	membrane fraction	*STMN2*
ENSG00000104435	GO:0048471	perinuclear region of cytoplasm	*STMN2*
ENSG00000104435	GO:0005794	Golgi apparatus	*STMN2*
ENSG00000104435	GO:0005768	endosome	*STMN2*
ENSG00000104435	GO:0030426	growth cone	*STMN2*
ENSG00000104435	GO:0005625	soluble fraction	*STMN2*
ENSG00000104435	GO:0031982	vesicle	*STMN2*
ENSG00000104435	GO:0005515	protein binding	*STMN2*
ENSG00000164458	GO:0007275	multicellular organismal development	*T*
ENSG00000164458	GO:0008284	positive regulation of cell proliferation	*T*
ENSG00000164458	GO:0007165	signal transduction	*T*
ENSG00000164458	GO:0006355	regulation of transcription, DNA-dependent	*T*
ENSG00000164458	GO:0048706	embryonic skeletal system development	*T*
ENSG00000164458	GO:0060070	canonical Wnt receptor signaling pathway	*T*
ENSG00000164458	GO:0003007	heart morphogenesis	*T*
ENSG00000164458	GO:0055007	cardiac muscle cell differentiation	*T*
ENSG00000164458	GO:0001843	neural tube closure	*T*

ENSG00000164458	GO:0001756	somitogenesis	*T*
ENSG00000164458	GO:0030903	notochord development	*T*
ENSG00000164458	GO:0023019	signal transduction involved in regulation of gene expression	*T*
ENSG00000164458	GO:0060395	SMAD protein signal transduction	*T*
ENSG00000164458	GO:0007341	penetration of zona pellucida	*T*
ENSG00000164458	GO:0030509	BMP signaling pathway	*T*
ENSG00000164458	GO:0001570	vasculogenesis	*T*
ENSG00000164458	GO:0007498	mesoderm development	*T*
ENSG00000164458	GO:0008595	anterior/posterior axis specification, embryo	*T*
ENSG00000164458	GO:0060349	bone morphogenesis	*T*
ENSG00000164458	GO:0035121	tail morphogenesis	*T*
ENSG00000164458	GO:0090009	primitive streak formation	*T*
ENSG00000164458	GO:0001839	neural plate morphogenesis	*T*
ENSG00000164458	GO:0003257	positive regulation of transcription from RNA polymerase II promoter involved in myocardial precursor cell differentiation	*T*
ENSG00000164458	GO:0007509	mesoderm migration involved in gastrulation	*T*
ENSG00000164458	GO:0022414	reproductive process	*T*
ENSG00000164458	GO:0061371	determination of heart left/right asymmetry	*T*
ENSG00000164458	GO:0009653	anatomical structure morphogenesis	*T*
ENSG00000164458	GO:0006366	transcription from RNA polymerase II promoter	*T*
ENSG00000164458	GO:0009952	anterior/posterior pattern specification	*T*
ENSG00000164458	GO:0006357	regulation of transcription from RNA polymerase II promoter	*T*
ENSG00000164458	GO:0045944	positive regulation of transcription from RNA polymerase II promoter	*T*
ENSG00000164458	GO:0005634	nucleus	*T*
ENSG00000164458	GO:0005737	cytoplasm	*T*
ENSG00000164458	GO:0000790	nuclear chromatin	*T*
ENSG00000164458	GO:0000785	chromatin	*T*
ENSG00000164458	GO:0003677	DNA binding	*T*
ENSG00000164458	GO:0003700	sequence-specific DNA binding transcription factor activity	*T*
ENSG00000164458	GO:0043565	sequence-specific DNA binding	*T*
ENSG00000164458	GO:0000980	RNA polymerase II distal enhancer sequence-specific DNA binding	*T*
ENSG00000164458	GO:0003705	sequence-specific distal enhancer binding RNA polymerase II transcription factor activity	*T*
ENSG00000164458	GO:0000981	sequence-specific DNA binding RNA polymerase II transcription factor activity	*T*
ENSG00000164458	GO:0005515	protein binding	*T*
ENSG00000135605	GO:0035556	intracellular signal transduction	*TEC*
ENSG00000135605	GO:0006468	protein phosphorylation	*TEC*
ENSG00000135605	GO:0007229	integrin-mediated signaling pathway	*TEC*
ENSG00000135605	GO:0007243	intracellular protein kinase cascade	*TEC*
ENSG00000135605	GO:0050853	B cell receptor signaling pathway	*TEC*
ENSG00000135605	GO:0010543	regulation of platelet activation	*TEC*
ENSG00000135605	GO:0005737	cytoplasm	*TEC*
ENSG00000135605	GO:0005886	plasma membrane	*TEC*
ENSG00000135605	GO:0005829	cytosol	*TEC*
ENSG00000135605	GO:0005856	cytoskeleton	*TEC*
ENSG00000135605	GO:0000166	nucleotide binding	*TEC*
ENSG00000135605	GO:0005515	protein binding	*TEC*
ENSG00000135605	GO:0005524	ATP binding	*TEC*

ENSG00000135605	GO:0005543	phospholipid binding	*TEC*
ENSG00000135605	GO:0046872	metal ion binding	*TEC*
ENSG00000135605	GO:0004672	protein kinase activity	*TEC*
ENSG00000135605	GO:0004713	protein tyrosine kinase activity	*TEC*
ENSG00000135605	GO:0004715	non-membrane spanning protein tyrosine kinase activity	*TEC*
ENSG00000159445	GO:0048015	phosphatidylinositol-mediated signaling	*THEM4*
ENSG00000159445	GO:0008286	insulin receptor signaling pathway	*THEM4*
ENSG00000159445	GO:0007173	epidermal growth factor receptor signaling pathway	*THEM4*
ENSG00000159445	GO:0008543	fibroblast growth factor receptor signaling pathway	*THEM4*
ENSG00000159445	GO:0048011	nerve growth factor receptor signaling pathway	*THEM4*
ENSG00000159445	GO:0005737	cytoplasm	*THEM4*
ENSG00000159445	GO:0005886	plasma membrane	*THEM4*
ENSG00000159445	GO:0042995	cell projection	*THEM4*
ENSG00000159445	GO:0005739	mitochondrion	*THEM4*
ENSG00000159445	GO:0005829	cytosol	*THEM4*
ENSG00000159445	GO:0043231	intracellular membrane-bounded organelle	*THEM4*
ENSG00000159445	GO:0032587	ruffle membrane	*THEM4*
ENSG00000139173	GO:0008150	biological_process	*TMEM117*
ENSG00000139173	GO:0016020	membrane	*TMEM117*
ENSG00000139173	GO:0016021	integral to membrane	*TMEM117*
ENSG00000139173	GO:0005783	endoplasmic reticulum	*TMEM117*
ENSG00000139173	GO:0003674	molecular_function	*TMEM117*
ENSG00000173452	GO:0016020	membrane	*TMEM196*
ENSG00000173452	GO:0016021	integral to membrane	*TMEM196*
ENSG00000106025	GO:0007166	cell surface receptor signaling pathway	*TSPAN12*
ENSG00000106025	GO:0045765	regulation of angiogenesis	*TSPAN12*
ENSG00000106025	GO:0010842	retina layer formation	*TSPAN12*
ENSG00000106025	GO:0016020	membrane	*TSPAN12*
ENSG00000106025	GO:0016021	integral to membrane	*TSPAN12*
ENSG00000106025	GO:0005886	plasma membrane	*TSPAN12*
ENSG00000106025	GO:0005624	membrane fraction	*TSPAN12*
ENSG00000106025	GO:0005887	integral to plasma membrane	*TSPAN12*
ENSG00000186153	GO:0055114	oxidation-reduction process	*WWOX*
ENSG00000186153	GO:0045944	positive regulation of transcription from RNA polymerase II promoter	*WWOX*
ENSG00000186153	GO:0008152	metabolic process	*WWOX*
ENSG00000186153	GO:0006917	induction of apoptosis	*WWOX*
ENSG00000186153	GO:0071560	cellular response to transforming growth factor beta stimulus	*WWOX*
ENSG00000186153	GO:0048705	skeletal system morphogenesis	*WWOX*
ENSG00000186153	GO:0043065	positive regulation of apoptotic process	*WWOX*
ENSG00000186153	GO:0030178	negative regulation of Wnt receptor signaling pathway	*WWOX*
ENSG00000186153	GO:0008202	steroid metabolic process	*WWOX*
ENSG00000186153	GO:0001649	osteoblast differentiation	*WWOX*
ENSG00000186153	GO:0005634	nucleus	*WWOX*
ENSG00000186153	GO:0005737	cytoplasm	*WWOX*
ENSG00000186153	GO:0005886	plasma membrane	*WWOX*
ENSG00000186153	GO:0005739	mitochondrion	*WWOX*
ENSG00000186153	GO:0005829	cytosol	*WWOX*
ENSG00000186153	GO:0005794	Golgi apparatus	*WWOX*
ENSG00000186153	GO:0005902	microvillus	*WWOX*

ENSG00000186153	GO:0000166	nucleotide binding	*WWOX*
ENSG00000186153	GO:0005515	protein binding	*WWOX*
ENSG00000186153	GO:0016491	oxidoreductase activity	*WWOX*
ENSG00000186153	GO:0050662	coenzyme binding	*WWOX*
ENSG00000186153	GO:0019899	enzyme binding	*WWOX*
ENSG00000186153	GO:0046983	protein dimerization activity	*WWOX*
ENSG00000186153	GO:0048037	cofactor binding	*WWOX*
ENSG00000108039	GO:0006508	proteolysis	*XPNPEP1*
ENSG00000108039	GO:0009987	cellular process	*XPNPEP1*
ENSG00000108039	GO:0010815	bradykinin catabolic process	*XPNPEP1*
ENSG00000108039	GO:0005737	cytoplasm	*XPNPEP1*
ENSG00000108039	GO:0005829	cytosol	*XPNPEP1*
ENSG00000108039	GO:0042803	protein homodimerization activity	*XPNPEP1*
ENSG00000108039	GO:0016787	hydrolase activity	*XPNPEP1*
ENSG00000108039	GO:0046872	metal ion binding	*XPNPEP1*
ENSG00000108039	GO:0008233	peptidase activity	*XPNPEP1*
ENSG00000108039	GO:0004177	aminopeptidase activity	*XPNPEP1*
ENSG00000108039	GO:0008237	metallopeptidase activity	*XPNPEP1*
ENSG00000108039	GO:0030145	manganese ion binding	*XPNPEP1*
ENSG00000108039	GO:0070006	metalloaminopeptidase activity	*XPNPEP1*
ENSG00000170396	GO:0005622	intracellular	*ZNF804A*
ENSG00000170396	GO:0008270	zinc ion binding	*ZNF804A*
ENSG00000170396	GO:0046872	metal ion binding	*ZNF804A*

Appendix III- WTCCC Phase II Ulcerative Colitis Cases, NBS Controls and 1958BC Controls- Quality Control and Data manipulation prior to analysis flowchart.

STEP 1- SENDING A REQUEST TO DOWNLOAD THE DATA & A REQUEST FOR ENCRYPTION KEYS TO BE SENT VIA POST

↓

STEP 2- DECRYPTION OF DATA

Download Gpg4Win from http://www.gnupg.org

↓

STEP 3- CONVERT FROM OXSTAT TO PLINK FORMAT

Download the static executable version of GTOOL from http://www.well.ox.ac.uk/~cfreeman/software/gwas/gtool.html. The GTOOL program will be fed the extracted CHIAMO.gen file (genotype calls) for each Chromosome and the sample file.

↓

STEP 4 CONVERTING 'NN' FOR '00' FOR MISSING GENOTYPES

↓

STEP 5- DATA FILTERING ACCORDING TO PROVIDED SNP EXCLUSION LISTS

STEP 5- PERFORMING QUALITY CONTROL USING R SNPRelate AND GDSFMT PACKAGEs

This step was carried out using the instructions here: http://corearray.sourceforge.net/tutorials/SNPRelate/
Analysis carried out: Principal Components Analysis, IBD Method of Moments, and Multi-Dimensional Scaling Analysis using IBS distances. QC filtering used for snps: MAF cut-off = 0.05, missing rate cut-off= 0.05 and ld.threshold = 0.2

↓

STEP 6- REMOVING OUTLIERS EMERGING FROM STEP 5

↓

STEP 6- CONVERT PLINK TO LDU MAPPING APPROACH FORMAT

↓

STEP 6- CONCATENATION OF CONTROL AND CASES DATASETS

↓

STEP 7- DATA GENOTYPE ENCODING (AA, AT…→ 11, 12, …)

Encoding all the genotypes (A A, A C…) into number codes (1 1, 1 2 …). This step also produces details about the MAF (details_genotype.txt) and Hardy Weinberg (fulldetails_genotype.txt- also for MAF)-

↓

STEP 8- DATA SNP FILTERING - IN HOUSE QC

Filtering of the SNPs which fail HWE test, have a low MAF (<0.03), HWE Chi Squared test >10 or have a percentage missingness > 20%. this information is provided in a file in the previous step (fulldetails_* file). This is only done on the id filtered concatenated controls dataset.

Appendix IV- NIDDK Ulcerative Colitis Cases and Crohn's Disease Controls- Quality Control and Data manipulation prior to analysis flowchart.

STEP 1- DOWNLOAD THE DATA

↓

STEP 2- DECRYPTION OF DATA

Using SRA Toolkit http://www.ncbi.nlm.nih.gov/Traces/sra/sra.cgi?view=toolkit_doc&f=std

↓

STEP 3 RECODING PLINK FILES IN PED AND MAP AND SPLITTING THESE GENOME WIDE FILES BY CHROMOSOME

↓

STEP 4 CONVERTING 'NN' FOR '00' FOR MISSING GENOTYPES AND ASSIGNING AFFECTION STATUS

↓

STEP 5- PERFORMING QUALITY CONTROL USING R SNPRelate AND GDSFMT PACKAGEs

This step was carried out using the instructions here: http://corearray.sourceforge.net/tutorials/SNPRelate/
Analysis carried out: Principal Components Analysis, IBD Method of Moments, and Multi-Dimensional Scaling Analysis using IBS distances. QC filtering used for snps: MAF cut-off = 0.05, missing rate cut-off= 0.05 and ld.threshold = 0.2 (see pages 22-66 of Lab book Volume 3)

↓

STEP 6- REMOVING OUTLIERS EMERGING FROM STEP 5

This is done by identifying the individual ID in the previous step and writing a perl program, which identifies the ID in the dataset and removes it- still to be optimised!

↓

STEP 7- CONVERT PLINK TO LDU MAPPING APPROACH FORMAT

↓

STEP 8- GENOTYPE ENCODING

Encoding all the genotypes (A A, A C…) into number codes (1 1, 1 2 …). This step also produces details about the MAF (details_genotype.txt) and Hardy Weinberg (fulldetails_genotype.txt- also for MAF)-

↓

STEP 9- - DATA SNP FILTERING -IN HOUSE QC

Filtering of the SNPs which fail HWE test, have a low MAF (<0.03), HWE Chi Squared test >10 or have a percentage missingness > 20%. this information is provided in a file in the previous step (fulldetails_* file). This is only done on the id filtered concatenated controls dataset.

Appendix V- BioMart Analysis from the Ulcerative Colitis GWAS described in Chapter IV

Ensembl Gene ID	GO Term Accession	GO Term Name	Associated Gene Name
ENSG00000006071	GO:0005524	ATP binding	*ABCC8*
ENSG00000006071	GO:0005886	plasma membrane	*ABCC8*
ENSG00000006071	GO:0005975	carbohydrate metabolic process	*ABCC8*
ENSG00000006071	GO:0006112	energy reserve metabolic process	*ABCC8*
ENSG00000006071	GO:0006200	ATP catabolic process	*ABCC8*
ENSG00000006071	GO:0006810	transport	*ABCC8*
ENSG00000006071	GO:0006813	potassium ion transport	*ABCC8*
ENSG00000006071	GO:0007165	signal transduction	*ABCC8*
ENSG00000006071	GO:0007268	synaptic transmission	*ABCC8*
ENSG00000006071	GO:0008076	voltage-gated potassium channel complex	*ABCC8*
ENSG00000006071	GO:0008281	sulfonylurea receptor activity	*ABCC8*
ENSG00000006071	GO:0015079	potassium ion transmembrane transporter activity	*ABCC8*
ENSG00000006071	GO:0016021	integral component of membrane	*ABCC8*
ENSG00000006071	GO:0016887	ATPase activity	*ABCC8*
ENSG00000006071	GO:0042626	ATPase activity, coupled to transmembrane movement of substances	*ABCC8*
ENSG00000006071	GO:0044281	small molecule metabolic process	*ABCC8*
ENSG00000006071	GO:0044325	ion channel binding	*ABCC8*
ENSG00000006071	GO:0050796	regulation of insulin secretion	*ABCC8*
ENSG00000006071	GO:0055085	transmembrane transport	*ABCC8*
ENSG00000006071	GO:0071805	potassium ion transmembrane transport	*ABCC8*
ENSG00000138075	GO:0005515	protein binding	*ABCG5*
ENSG00000138075	GO:0005524	ATP binding	*ABCG5*
ENSG00000138075	GO:0005886	plasma membrane	*ABCG5*
ENSG00000138075	GO:0006200	ATP catabolic process	*ABCG5*
ENSG00000138075	GO:0007584	response to nutrient	*ABCG5*
ENSG00000138075	GO:0007588	excretion	*ABCG5*
ENSG00000138075	GO:0010212	response to ionizing radiation	*ABCG5*
ENSG00000138075	GO:0010949	negative regulation of intestinal phytosterol absorption	*ABCG5*
ENSG00000138075	GO:0015918	sterol transport	*ABCG5*
ENSG00000138075	GO:0016020	membrane	*ABCG5*
ENSG00000138075	GO:0016021	integral component of membrane	*ABCG5*
ENSG00000138075	GO:0016324	apical plasma membrane	*ABCG5*
ENSG00000138075	GO:0016887	ATPase activity	*ABCG5*
ENSG00000138075	GO:0017127	cholesterol transporter activity	*ABCG5*
ENSG00000138075	GO:0030299	intestinal cholesterol absorption	*ABCG5*
ENSG00000138075	GO:0033344	cholesterol efflux	*ABCG5*
ENSG00000138075	GO:0042493	response to drug	*ABCG5*
ENSG00000138075	GO:0042632	cholesterol homeostasis	*ABCG5*
ENSG00000138075	GO:0044281	small molecule metabolic process	*ABCG5*
ENSG00000138075	GO:0045177	apical part of cell	*ABCG5*
ENSG00000138075	GO:0045796	negative regulation of intestinal cholesterol absorption	*ABCG5*
ENSG00000138075	GO:0046982	protein heterodimerization activity	*ABCG5*

ENSG00000138075	GO:0055085	transmembrane transport	*ABCG5*
ENSG00000107796	GO:0005515	protein binding	*ACTA2*
ENSG00000107796	GO:0005524	ATP binding	*ACTA2*
ENSG00000107796	GO:0005615	extracellular space	*ACTA2*
ENSG00000107796	GO:0005737	cytoplasm	*ACTA2*
ENSG00000107796	GO:0005829	cytosol	*ACTA2*
ENSG00000107796	GO:0005856	cytoskeleton	*ACTA2*
ENSG00000107796	GO:0006936	muscle contraction	*ACTA2*
ENSG00000107796	GO:0008217	regulation of blood pressure	*ACTA2*
ENSG00000107796	GO:0009615	response to virus	*ACTA2*
ENSG00000107796	GO:0014829	vascular smooth muscle contraction	*ACTA2*
ENSG00000107796	GO:0015629	actin cytoskeleton	*ACTA2*
ENSG00000107796	GO:0019901	protein kinase binding	*ACTA2*
ENSG00000107796	GO:0030485	smooth muscle contractile fiber	*ACTA2*
ENSG00000107796	GO:0043234	protein complex	*ACTA2*
ENSG00000107796	GO:0072144	glomerular mesangial cell development	*ACTA2*
ENSG00000157985	GO:0005515	protein binding	*AGAP1*
ENSG00000157985	GO:0005525	GTP binding	*AGAP1*
ENSG00000157985	GO:0005543	phospholipid binding	*AGAP1*
ENSG00000157985	GO:0005622	intracellular	*AGAP1*
ENSG00000157985	GO:0005737	cytoplasm	*AGAP1*
ENSG00000157985	GO:0006184	GTP catabolic process	*AGAP1*
ENSG00000157985	GO:0007165	signal transduction	*AGAP1*
ENSG00000157985	GO:0007264	small GTPase mediated signal transduction	*AGAP1*
ENSG00000157985	GO:0008060	ARF GTPase activator activity	*AGAP1*
ENSG00000157985	GO:0008270	zinc ion binding	*AGAP1*
ENSG00000157985	GO:0015031	protein transport	*AGAP1*
ENSG00000157985	GO:0016020	membrane	*AGAP1*
ENSG00000157985	GO:0032312	regulation of ARF GTPase activity	*AGAP1*
ENSG00000157985	GO:0043547	positive regulation of GTPase activity	*AGAP1*
ENSG00000171094	GO:0000187	activation of MAPK activity	*ALK*
ENSG00000171094	GO:0004672	protein kinase activity	*ALK*
ENSG00000171094	GO:0004704	NF-kappaB-inducing kinase activity	*ALK*
ENSG00000171094	GO:0004713	protein tyrosine kinase activity	*ALK*
ENSG00000171094	GO:0004714	transmembrane receptor protein tyrosine kinase activity	*ALK*
ENSG00000171094	GO:0005515	protein binding	*ALK*
ENSG00000171094	GO:0005524	ATP binding	*ALK*
ENSG00000171094	GO:0005887	integral component of plasma membrane	*ALK*
ENSG00000171094	GO:0006468	protein phosphorylation	*ALK*
ENSG00000171094	GO:0007165	signal transduction	*ALK*
ENSG00000171094	GO:0007169	transmembrane receptor protein tyrosine kinase signaling pathway	*ALK*
ENSG00000171094	GO:0007399	nervous system development	*ALK*
ENSG00000171094	GO:0008283	cell proliferation	*ALK*
ENSG00000171094	GO:0016020	membrane	*ALK*
ENSG00000171094	GO:0016310	phosphorylation	*ALK*
ENSG00000171094	GO:0016772	transferase activity, transferring phosphorus-containing groups	*ALK*
ENSG00000171094	GO:0018108	peptidyl-tyrosine phosphorylation	*ALK*
ENSG00000171094	GO:0038061	NIK/NF-kappaB signaling	*ALK*
ENSG00000171094	GO:0042981	regulation of apoptotic process	*ALK*
ENSG00000171094	GO:0043234	protein complex	*ALK*

ENSG00000171094	GO:0046777	protein autophosphorylation	*ALK*
ENSG00000171094	GO:0048666	neuron development	*ALK*
ENSG00000171094	GO:0051092	positive regulation of NF-kappaB transcription factor activity	*ALK*
ENSG00000171094	GO:0070062	extracellular vesicular exosome	*ALK*
ENSG00000148513	GO:0003677	DNA binding	*ANKRD30A*
ENSG00000148513	GO:0003700	sequence-specific DNA binding transcription factor activity	*ANKRD30A*
ENSG00000148513	GO:0005515	protein binding	*ANKRD30A*
ENSG00000148513	GO:0005634	nucleus	*ANKRD30A*
ENSG00000148513	GO:0006355	regulation of transcription, DNA-templated	*ANKRD30A*
ENSG00000071205	GO:0005100	Rho GTPase activator activity	*ARHGAP10*
ENSG00000071205	GO:0005515	protein binding	*ARHGAP10*
ENSG00000071205	GO:0005622	intracellular	*ARHGAP10*
ENSG00000071205	GO:0005829	cytosol	*ARHGAP10*
ENSG00000071205	GO:0005886	plasma membrane	*ARHGAP10*
ENSG00000071205	GO:0006915	apoptotic process	*ARHGAP10*
ENSG00000071205	GO:0007010	cytoskeleton organization	*ARHGAP10*
ENSG00000071205	GO:0007165	signal transduction	*ARHGAP10*
ENSG00000071205	GO:0007264	small GTPase mediated signal transduction	*ARHGAP10*
ENSG00000071205	GO:0032321	positive regulation of Rho GTPase activity	*ARHGAP10*
ENSG00000071205	GO:0042981	regulation of apoptotic process	*ARHGAP10*
ENSG00000071205	GO:0043066	negative regulation of apoptotic process	*ARHGAP10*
ENSG00000071205	GO:0048471	perinuclear region of cytoplasm	*ARHGAP10*
ENSG00000071205	GO:0051056	regulation of small GTPase mediated signal transduction	*ARHGAP10*
ENSG00000133794	GO:0000060	protein import into nucleus, translocation	*ARNTL*
ENSG00000133794	GO:0000976	transcription regulatory region sequence-specific DNA binding	*ARNTL*
ENSG00000133794	GO:0000982	RNA polymerase II core promoter proximal region sequence-specific DNA binding transcription factor activity	*ARNTL*
ENSG00000133794	GO:0001047	core promoter binding	*ARNTL*
ENSG00000133794	GO:0001190	RNA polymerase II transcription factor binding transcription factor activity involved in positive regulation of transcription	*ARNTL*
ENSG00000133794	GO:0003677	DNA binding	*ARNTL*
ENSG00000133794	GO:0003700	sequence-specific DNA binding transcription factor activity	*ARNTL*
ENSG00000133794	GO:0004871	signal transducer activity	*ARNTL*
ENSG00000133794	GO:0005515	protein binding	*ARNTL*
ENSG00000133794	GO:0005634	nucleus	*ARNTL*
ENSG00000133794	GO:0005667	transcription factor complex	*ARNTL*
ENSG00000133794	GO:0005737	cytoplasm	*ARNTL*
ENSG00000133794	GO:0006355	regulation of transcription, DNA-templated	*ARNTL*
ENSG00000133794	GO:0006366	transcription from RNA polymerase II promoter	*ARNTL*
ENSG00000133794	GO:0007165	signal transduction	*ARNTL*
ENSG00000133794	GO:0007283	spermatogenesis	*ARNTL*
ENSG00000133794	GO:0007623	circadian rhythm	*ARNTL*

ENSG00000133794	GO:0016604	nuclear body	*ARNTL*
ENSG00000133794	GO:0016605	PML body	*ARNTL*
ENSG00000133794	GO:0017162	aryl hydrocarbon receptor binding	*ARNTL*
ENSG00000133794	GO:0032007	negative regulation of TOR signaling	*ARNTL*
ENSG00000133794	GO:0032922	circadian regulation of gene expression	*ARNTL*
ENSG00000133794	GO:0033391	chromatoid body	*ARNTL*
ENSG00000133794	GO:0042176	regulation of protein catabolic process	*ARNTL*
ENSG00000133794	GO:0042634	regulation of hair cycle	*ARNTL*
ENSG00000133794	GO:0042753	positive regulation of circadian rhythm	*ARNTL*
ENSG00000133794	GO:0043161	proteasome-mediated ubiquitin-dependent protein catabolic process	*ARNTL*
ENSG00000133794	GO:0043425	bHLH transcription factor binding	*ARNTL*
ENSG00000133794	GO:0043565	sequence-specific DNA binding	*ARNTL*
ENSG00000133794	GO:0045599	negative regulation of fat cell differentiation	*ARNTL*
ENSG00000133794	GO:0045892	negative regulation of transcription, DNA-templated	*ARNTL*
ENSG00000133794	GO:0045893	positive regulation of transcription, DNA-templated	*ARNTL*
ENSG00000133794	GO:0045944	positive regulation of transcription from RNA polymerase II promoter	*ARNTL*
ENSG00000133794	GO:0046982	protein heterodimerization activity	*ARNTL*
ENSG00000133794	GO:0046983	protein dimerization activity	*ARNTL*
ENSG00000133794	GO:0050767	regulation of neurogenesis	*ARNTL*
ENSG00000133794	GO:0050796	regulation of insulin secretion	*ARNTL*
ENSG00000133794	GO:0051726	regulation of cell cycle	*ARNTL*
ENSG00000133794	GO:0051775	response to redox state	*ARNTL*
ENSG00000133794	GO:0051879	Hsp90 protein binding	*ARNTL*
ENSG00000133794	GO:0070888	E-box binding	*ARNTL*
ENSG00000133794	GO:0090263	positive regulation of canonical Wnt signaling pathway	*ARNTL*
ENSG00000133794	GO:0090403	oxidative stress-induced premature senescence	*ARNTL*
ENSG00000133794	GO:2000074	regulation of type B pancreatic cell development	*ARNTL*
ENSG00000133794	GO:2000323	negative regulation of glucocorticoid receptor signaling pathway	*ARNTL*
ENSG00000133794	GO:2000772	regulation of cellular senescence	*ARNTL*
ENSG00000133794	GO:2001016	positive regulation of skeletal muscle cell differentiation	*ARNTL*
ENSG00000172318	GO:0000139	Golgi membrane	*B3GALT1*
ENSG00000172318	GO:0006486	protein glycosylation	*B3GALT1*
ENSG00000172318	GO:0008378	galactosyltransferase activity	*B3GALT1*
ENSG00000172318	GO:0008499	UDP-galactose:beta-N-acetylglucosamine beta-1,3-galactosyltransferase activity	*B3GALT1*
ENSG00000172318	GO:0009312	oligosaccharide biosynthetic process	*B3GALT1*
ENSG00000172318	GO:0016020	membrane	*B3GALT1*
ENSG00000172318	GO:0016021	integral component of membrane	*B3GALT1*
ENSG00000161267	GO:0001889	liver development	*BDH1*
ENSG00000275544	GO:0003858	3-hydroxybutyrate dehydrogenase activity	*BDH1*
ENSG00000161267	GO:0003858	3-hydroxybutyrate dehydrogenase activity	*BDH1*
ENSG00000161267	GO:0005543	phospholipid binding	*BDH1*
ENSG00000275544	GO:0005634	nucleus	*BDH1*

ENSG00000161267	GO:0005634	nucleus	*BDH1*
ENSG00000275544	GO:0005739	mitochondrion	*BDH1*
ENSG00000161267	GO:0005739	mitochondrion	*BDH1*
ENSG00000161267	GO:0005743	mitochondrial inner membrane	*BDH1*
ENSG00000275544	GO:0005759	mitochondrial matrix	*BDH1*
ENSG00000161267	GO:0005759	mitochondrial matrix	*BDH1*
ENSG00000161267	GO:0007420	brain development	*BDH1*
ENSG00000161267	GO:0007584	response to nutrient	*BDH1*
ENSG00000275544	GO:0008152	metabolic process	*BDH1*
ENSG00000161267	GO:0008152	metabolic process	*BDH1*
ENSG00000161267	GO:0009636	response to toxic substance	*BDH1*
ENSG00000161267	GO:0009725	response to hormone	*BDH1*
ENSG00000275544	GO:0016491	oxidoreductase activity	*BDH1*
ENSG00000161267	GO:0016491	oxidoreductase activity	*BDH1*
ENSG00000161267	GO:0032355	response to estradiol	*BDH1*
ENSG00000161267	GO:0032868	response to insulin	*BDH1*
ENSG00000161267	GO:0042493	response to drug	*BDH1*
ENSG00000161267	GO:0042594	response to starvation	*BDH1*
ENSG00000275544	GO:0044255	cellular lipid metabolic process	*BDH1*
ENSG00000161267	GO:0044255	cellular lipid metabolic process	*BDH1*
ENSG00000275544	GO:0044281	small molecule metabolic process	*BDH1*
ENSG00000161267	GO:0044281	small molecule metabolic process	*BDH1*
ENSG00000161267	GO:0045471	response to ethanol	*BDH1*
ENSG00000161267	GO:0046686	response to cadmium ion	*BDH1*
ENSG00000275544	GO:0046950	cellular ketone body metabolic process	*BDH1*
ENSG00000161267	GO:0046950	cellular ketone body metabolic process	*BDH1*
ENSG00000275544	GO:0046951	ketone body biosynthetic process	*BDH1*
ENSG00000161267	GO:0046951	ketone body biosynthetic process	*BDH1*
ENSG00000275544	GO:0046952	ketone body catabolic process	*BDH1*
ENSG00000161267	GO:0046952	ketone body catabolic process	*BDH1*
ENSG00000161267	GO:0051412	response to corticosterone	*BDH1*
ENSG00000275544	GO:0055114	oxidation-reduction process	*BDH1*
ENSG00000161267	GO:0055114	oxidation-reduction process	*BDH1*
ENSG00000161267	GO:0060416	response to growth hormone	*BDH1*
ENSG00000161267	GO:0060612	adipose tissue development	*BDH1*
ENSG00000186190	GO:0005576	extracellular region	*BPIFB3*
ENSG00000186190	GO:0005737	cytoplasm	*BPIFB3*
ENSG00000186190	GO:0008289	lipid binding	*BPIFB3*
ENSG00000186190	GO:0045087	innate immune response	*BPIFB3*
ENSG00000186191	GO:0005576	extracellular region	*BPIFB4*
ENSG00000186191	GO:0005737	cytoplasm	*BPIFB4*
ENSG00000186191	GO:0008289	lipid binding	*BPIFB4*
ENSG00000136261	GO:0005488	binding	*BZW2*
ENSG00000136261	GO:0005515	protein binding	*BZW2*
ENSG00000136261	GO:0007399	nervous system development	*BZW2*
ENSG00000136261	GO:0016020	membrane	*BZW2*
ENSG00000136261	GO:0030154	cell differentiation	*BZW2*
ENSG00000120055	GO:0005515	protein binding	*C10orf95*
ENSG00000186493	GO:0005576	extracellular region	*C5orf38*
ENSG00000205174	GO:0016021	integral component of membrane	*C7orf66*
ENSG00000188782	GO:0001669	acrosomal vesicle	*CATSPER4*
ENSG00000188782	GO:0005216	ion channel activity	*CATSPER4*
ENSG00000188782	GO:0005227	calcium activated cation channel activity	*CATSPER4*

ENSG00000188782	GO:0005245	voltage-gated calcium channel activity	*CATSPER4*
ENSG00000188782	GO:0005515	protein binding	*CATSPER4*
ENSG00000188782	GO:0005886	plasma membrane	*CATSPER4*
ENSG00000188782	GO:0006811	ion transport	*CATSPER4*
ENSG00000188782	GO:0006812	cation transport	*CATSPER4*
ENSG00000188782	GO:0006814	sodium ion transport	*CATSPER4*
ENSG00000188782	GO:0007275	multicellular organismal development	*CATSPER4*
ENSG00000188782	GO:0007283	spermatogenesis	*CATSPER4*
ENSG00000188782	GO:0007338	single fertilization	*CATSPER4*
ENSG00000188782	GO:0016020	membrane	*CATSPER4*
ENSG00000188782	GO:0030154	cell differentiation	*CATSPER4*
ENSG00000188782	GO:0030317	sperm motility	*CATSPER4*
ENSG00000188782	GO:0032504	multicellular organism reproduction	*CATSPER4*
ENSG00000188782	GO:0034220	ion transmembrane transport	*CATSPER4*
ENSG00000188782	GO:0034765	regulation of ion transmembrane transport	*CATSPER4*
ENSG00000188782	GO:0035036	sperm-egg recognition	*CATSPER4*
ENSG00000188782	GO:0036128	CatSper complex	*CATSPER4*
ENSG00000188782	GO:0048240	sperm capacitation	*CATSPER4*
ENSG00000188782	GO:0055085	transmembrane transport	*CATSPER4*
ENSG00000188782	GO:0070509	calcium ion import	*CATSPER4*
ENSG00000188782	GO:0086010	membrane depolarization during action potential	*CATSPER4*
ENSG00000188782	GO:0097228	sperm principal piece	*CATSPER4*
ENSG00000054803	GO:0005515	protein binding	*CBLN4*
ENSG00000054803	GO:0005615	extracellular space	*CBLN4*
ENSG00000054803	GO:0009306	protein secretion	*CBLN4*
ENSG00000054803	GO:0030054	cell junction	*CBLN4*
ENSG00000054803	GO:0045202	synapse	*CBLN4*
ENSG00000123106	GO:0005634	nucleus	*CCDC91*
ENSG00000123106	GO:0005794	Golgi apparatus	*CCDC91*
ENSG00000123106	GO:0015031	protein transport	*CCDC91*
ENSG00000123106	GO:0016020	membrane	*CCDC91*
ENSG00000123106	GO:0042802	identical protein binding	*CCDC91*
ENSG00000111860	GO:0005737	cytoplasm	*CEP85L*
ENSG00000111860	GO:0005813	centrosome	*CEP85L*
ENSG00000174600	GO:0001501	skeletal system development	*CMKLR1*
ENSG00000174600	GO:0004872	receptor activity	*CMKLR1*
ENSG00000174600	GO:0004930	G-protein coupled receptor activity	*CMKLR1*
ENSG00000174600	GO:0004950	chemokine receptor activity	*CMKLR1*
ENSG00000174600	GO:0005515	protein binding	*CMKLR1*
ENSG00000174600	GO:0005886	plasma membrane	*CMKLR1*
ENSG00000174600	GO:0005887	integral component of plasma membrane	*CMKLR1*
ENSG00000174600	GO:0006935	chemotaxis	*CMKLR1*
ENSG00000174600	GO:0006955	immune response	*CMKLR1*
ENSG00000174600	GO:0007186	G-protein coupled receptor signaling pathway	*CMKLR1*
ENSG00000174600	GO:0010759	positive regulation of macrophage chemotaxis	*CMKLR1*
ENSG00000174600	GO:0016021	integral component of membrane	*CMKLR1*
ENSG00000174600	GO:0032088	negative regulation of NF-kappaB transcription factor activity	*CMKLR1*
ENSG00000174600	GO:0032695	negative regulation of interleukin-12 production	*CMKLR1*
ENSG00000174600	GO:0045600	positive regulation of fat cell	*CMKLR1*

		differentiation	
ENSG00000174600	GO:0050848	regulation of calcium-mediated signaling	*CMKLR1*
ENSG00000174600	GO:0070098	chemokine-mediated signaling pathway	*CMKLR1*
ENSG00000174469	GO:0005515	protein binding	*CNTNAP2*
ENSG00000174469	GO:0005769	early endosome	*CNTNAP2*
ENSG00000174469	GO:0005794	Golgi apparatus	*CNTNAP2*
ENSG00000174469	GO:0007155	cell adhesion	*CNTNAP2*
ENSG00000174469	GO:0007420	brain development	*CNTNAP2*
ENSG00000174469	GO:0007612	learning	*CNTNAP2*
ENSG00000174469	GO:0008038	neuron recognition	*CNTNAP2*
ENSG00000174469	GO:0008076	voltage-gated potassium channel complex	*CNTNAP2*
ENSG00000174469	GO:0009986	cell surface	*CNTNAP2*
ENSG00000174469	GO:0016020	membrane	*CNTNAP2*
ENSG00000174469	GO:0016021	integral component of membrane	*CNTNAP2*
ENSG00000174469	GO:0019226	transmission of nerve impulse	*CNTNAP2*
ENSG00000174469	GO:0019899	enzyme binding	*CNTNAP2*
ENSG00000174469	GO:0021756	striatum development	*CNTNAP2*
ENSG00000174469	GO:0021761	limbic system development	*CNTNAP2*
ENSG00000174469	GO:0021794	thalamus development	*CNTNAP2*
ENSG00000174469	GO:0021987	cerebral cortex development	*CNTNAP2*
ENSG00000174469	GO:0030424	axon	*CNTNAP2*
ENSG00000174469	GO:0030425	dendrite	*CNTNAP2*
ENSG00000174469	GO:0030534	adult behavior	*CNTNAP2*
ENSG00000174469	GO:0030673	axolemma	*CNTNAP2*
ENSG00000174469	GO:0031175	neuron projection development	*CNTNAP2*
ENSG00000174469	GO:0034613	cellular protein localization	*CNTNAP2*
ENSG00000174469	GO:0035176	social behavior	*CNTNAP2*
ENSG00000174469	GO:0043025	neuronal cell body	*CNTNAP2*
ENSG00000174469	GO:0043204	perikaryon	*CNTNAP2*
ENSG00000174469	GO:0044224	juxtaparanode region of axon	*CNTNAP2*
ENSG00000174469	GO:0045163	clustering of voltage-gated potassium channels	*CNTNAP2*
ENSG00000174469	GO:0071109	superior temporal gyrus development	*CNTNAP2*
ENSG00000174469	GO:0071205	protein localization to juxtaparanode region of axon	*CNTNAP2*
ENSG00000174469	GO:0071625	vocalization behavior	*CNTNAP2*
ENSG00000133103	GO:0000139	Golgi membrane	*COG6*
ENSG00000133103	GO:0006891	intra-Golgi vesicle-mediated transport	*COG6*
ENSG00000133103	GO:0015031	protein transport	*COG6*
ENSG00000133103	GO:0017119	Golgi transport complex	*COG6*
ENSG00000133103	GO:0070085	glycosylation	*COG6*
ENSG00000145781	GO:0005515	protein binding	*COMMD10*
ENSG00000173085	GO:0002083	4-hydroxybenzoate decaprenyltransferase activity	*COQ2*
ENSG00000173085	GO:0004659	prenyltransferase activity	*COQ2*
ENSG00000173085	GO:0005739	mitochondrion	*COQ2*
ENSG00000173085	GO:0005743	mitochondrial inner membrane	*COQ2*
ENSG00000173085	GO:0006071	glycerol metabolic process	*COQ2*
ENSG00000173085	GO:0006744	ubiquinone biosynthetic process	*COQ2*
ENSG00000173085	GO:0008152	metabolic process	*COQ2*
ENSG00000173085	GO:0008219	cell death	*COQ2*
ENSG00000173085	GO:0008299	isoprenoid biosynthetic process	*COQ2*
ENSG00000173085	GO:0009058	biosynthetic process	*COQ2*

ENSG00000173085	GO:0016021	integral component of membrane	*COQ2*
ENSG00000173085	GO:0044281	small molecule metabolic process	*COQ2*
ENSG00000173085	GO:0047293	4-hydroxybenzoate nonaprenyltransferase activity	*COQ2*
ENSG00000134030	GO:0000184	nuclear-transcribed mRNA catabolic process, nonsense-mediated decay	*CTIF*
ENSG00000134030	GO:0003723	RNA binding	*CTIF*
ENSG00000134030	GO:0005488	binding	*CTIF*
ENSG00000134030	GO:0005515	protein binding	*CTIF*
ENSG00000134030	GO:0005737	cytoplasm	*CTIF*
ENSG00000134030	GO:0006446	regulation of translational initiation	*CTIF*
ENSG00000134030	GO:0048471	perinuclear region of cytoplasm	*CTIF*
ENSG00000066032	GO:0005198	structural molecule activity	*CTNNA2*
ENSG00000066032	GO:0005200	structural constituent of cytoskeleton	*CTNNA2*
ENSG00000066032	GO:0005515	protein binding	*CTNNA2*
ENSG00000066032	GO:0005737	cytoplasm	*CTNNA2*
ENSG00000066032	GO:0005829	cytosol	*CTNNA2*
ENSG00000066032	GO:0005912	adherens junction	*CTNNA2*
ENSG00000066032	GO:0005913	cell-cell adherens junction	*CTNNA2*
ENSG00000066032	GO:0007155	cell adhesion	*CTNNA2*
ENSG00000066032	GO:0007409	axonogenesis	*CTNNA2*
ENSG00000066032	GO:0015629	actin cytoskeleton	*CTNNA2*
ENSG00000066032	GO:0016323	basolateral plasma membrane	*CTNNA2*
ENSG00000066032	GO:0016337	single organismal cell-cell adhesion	*CTNNA2*
ENSG00000066032	GO:0021942	radial glia guided migration of Purkinje cell	*CTNNA2*
ENSG00000066032	GO:0030027	lamellipodium	*CTNNA2*
ENSG00000066032	GO:0030424	axon	*CTNNA2*
ENSG00000066032	GO:0042692	muscle cell differentiation	*CTNNA2*
ENSG00000066032	GO:0045296	cadherin binding	*CTNNA2*
ENSG00000066032	GO:0048813	dendrite morphogenesis	*CTNNA2*
ENSG00000066032	GO:0048854	brain morphogenesis	*CTNNA2*
ENSG00000066032	GO:0051015	actin filament binding	*CTNNA2*
ENSG00000066032	GO:0051149	positive regulation of muscle cell differentiation	*CTNNA2*
ENSG00000066032	GO:0051823	regulation of synapse structural plasticity	*CTNNA2*
ENSG00000066032	GO:0060134	prepulse inhibition	*CTNNA2*
ENSG00000121966	GO:0000187	activation of MAPK activity	*CXCR4*
ENSG00000121966	GO:0001569	patterning of blood vessels	*CXCR4*
ENSG00000121966	GO:0001618	virus receptor activity	*CXCR4*
ENSG00000121966	GO:0001666	response to hypoxia	*CXCR4*
ENSG00000121966	GO:0001667	ameboidal cell migration	*CXCR4*
ENSG00000121966	GO:0001764	neuron migration	*CXCR4*
ENSG00000121966	GO:0002407	dendritic cell chemotaxis	*CXCR4*
ENSG00000121966	GO:0003779	actin binding	*CXCR4*
ENSG00000121966	GO:0004930	G-protein coupled receptor activity	*CXCR4*
ENSG00000121966	GO:0005515	protein binding	*CXCR4*
ENSG00000121966	GO:0005737	cytoplasm	*CXCR4*
ENSG00000121966	GO:0005764	lysosome	*CXCR4*
ENSG00000121966	GO:0005769	early endosome	*CXCR4*
ENSG00000121966	GO:0005770	late endosome	*CXCR4*
ENSG00000121966	GO:0005886	plasma membrane	*CXCR4*
ENSG00000121966	GO:0005911	cell-cell junction	*CXCR4*
ENSG00000121966	GO:0006915	apoptotic process	*CXCR4*
ENSG00000121966	GO:0006954	inflammatory response	*CXCR4*

ENSG00000121966	GO:0007186	G-protein coupled receptor signaling pathway	*CXCR4*
ENSG00000121966	GO:0007204	positive regulation of cytosolic calcium ion concentration	*CXCR4*
ENSG00000121966	GO:0007281	germ cell development	*CXCR4*
ENSG00000121966	GO:0007420	brain development	*CXCR4*
ENSG00000121966	GO:0008045	motor neuron axon guidance	*CXCR4*
ENSG00000121966	GO:0008354	germ cell migration	*CXCR4*
ENSG00000121966	GO:0009615	response to virus	*CXCR4*
ENSG00000121966	GO:0009897	external side of plasma membrane	*CXCR4*
ENSG00000121966	GO:0009986	cell surface	*CXCR4*
ENSG00000121966	GO:0015026	coreceptor activity	*CXCR4*
ENSG00000121966	GO:0016021	integral component of membrane	*CXCR4*
ENSG00000121966	GO:0016023	cytoplasmic membrane-bounded vesicle	*CXCR4*
ENSG00000121966	GO:0016032	viral process	*CXCR4*
ENSG00000121966	GO:0016494	C-X-C chemokine receptor activity	*CXCR4*
ENSG00000121966	GO:0019722	calcium-mediated signaling	*CXCR4*
ENSG00000121966	GO:0019955	cytokine binding	*CXCR4*
ENSG00000121966	GO:0030054	cell junction	*CXCR4*
ENSG00000121966	GO:0030260	entry into host cell	*CXCR4*
ENSG00000121966	GO:0030334	regulation of cell migration	*CXCR4*
ENSG00000121966	GO:0030426	growth cone	*CXCR4*
ENSG00000121966	GO:0031252	cell leading edge	*CXCR4*
ENSG00000121966	GO:0031410	cytoplasmic vesicle	*CXCR4*
ENSG00000121966	GO:0031625	ubiquitin protein ligase binding	*CXCR4*
ENSG00000121966	GO:0032027	myosin light chain binding	*CXCR4*
ENSG00000121966	GO:0042098	T cell proliferation	*CXCR4*
ENSG00000121966	GO:0042119	neutrophil activation	*CXCR4*
ENSG00000121966	GO:0043130	ubiquitin binding	*CXCR4*
ENSG00000121966	GO:0043217	myelin maintenance	*CXCR4*
ENSG00000121966	GO:0048699	generation of neurons	*CXCR4*
ENSG00000121966	GO:0048714	positive regulation of oligodendrocyte differentiation	*CXCR4*
ENSG00000121966	GO:0050920	regulation of chemotaxis	*CXCR4*
ENSG00000121966	GO:0061351	neural precursor cell proliferation	*CXCR4*
ENSG00000121966	GO:0070062	extracellular vesicular exosome	*CXCR4*
ENSG00000121966	GO:0070098	chemokine-mediated signaling pathway	*CXCR4*
ENSG00000121966	GO:0071345	cellular response to cytokine stimulus	*CXCR4*
ENSG00000164821	GO:0005576	extracellular region	*DEFA4*
ENSG00000164821	GO:0005615	extracellular space	*DEFA4*
ENSG00000164821	GO:0005796	Golgi lumen	*DEFA4*
ENSG00000164821	GO:0006952	defense response	*DEFA4*
ENSG00000164821	GO:0019731	antibacterial humoral response	*DEFA4*
ENSG00000164821	GO:0031640	killing of cells of other organism	*DEFA4*
ENSG00000164821	GO:0045087	innate immune response	*DEFA4*
ENSG00000164821	GO:0050832	defense response to fungus	*DEFA4*
ENSG00000164825	GO:0002227	innate immune response in mucosa	*DEFB1*
ENSG00000164825	GO:0002526	acute inflammatory response	*DEFB1*
ENSG00000164825	GO:0005576	extracellular region	*DEFB1*
ENSG00000164825	GO:0005615	extracellular space	*DEFB1*
ENSG00000164825	GO:0005796	Golgi lumen	*DEFB1*
ENSG00000164825	GO:0006935	chemotaxis	*DEFB1*
ENSG00000164825	GO:0006952	defense response	*DEFB1*
ENSG00000164825	GO:0006955	immune response	*DEFB1*

ENSG00000164825	GO:0007186	G-protein coupled receptor signaling pathway	*DEFB1*
ENSG00000164825	GO:0009617	response to bacterium	*DEFB1*
ENSG00000164825	GO:0019731	antibacterial humoral response	*DEFB1*
ENSG00000164825	GO:0033574	response to testosterone	*DEFB1*
ENSG00000164825	GO:0045087	innate immune response	*DEFB1*
ENSG00000164825	GO:0050830	defense response to Gram-positive bacterium	*DEFB1*
ENSG00000164825	GO:0070062	extracellular vesicular exosome	*DEFB1*
ENSG00000173253	GO:0003674	molecular_function	*DMRT2*
ENSG00000173253	GO:0003700	sequence-specific DNA binding transcription factor activity	*DMRT2*
ENSG00000173253	GO:0005575	cellular_component	*DMRT2*
ENSG00000173253	GO:0005634	nucleus	*DMRT2*
ENSG00000173253	GO:0006351	transcription, DNA-templated	*DMRT2*
ENSG00000173253	GO:0006355	regulation of transcription, DNA-templated	*DMRT2*
ENSG00000173253	GO:0007530	sex determination	*DMRT2*
ENSG00000173253	GO:0008150	biological_process	*DMRT2*
ENSG00000173253	GO:0008584	male gonad development	*DMRT2*
ENSG00000173253	GO:0014807	regulation of somitogenesis	*DMRT2*
ENSG00000173253	GO:0042803	protein homodimerization activity	*DMRT2*
ENSG00000173253	GO:0043565	sequence-specific DNA binding	*DMRT2*
ENSG00000173253	GO:0046872	metal ion binding	*DMRT2*
ENSG00000173253	GO:0048706	embryonic skeletal system development	*DMRT2*
ENSG00000173253	GO:2000287	positive regulation of myotome development	*DMRT2*
ENSG00000206052	GO:0005158	insulin receptor binding	*DOK6*
ENSG00000206052	GO:0005515	protein binding	*DOK6*
ENSG00000117543	GO:0004164	diphthine synthase activity	*DPH5*
ENSG00000117543	GO:0005575	cellular_component	*DPH5*
ENSG00000117543	GO:0008152	metabolic process	*DPH5*
ENSG00000117543	GO:0008168	methyltransferase activity	*DPH5*
ENSG00000117543	GO:0017183	peptidyl-diphthamide biosynthetic process from peptidyl-histidine	*DPH5*
ENSG00000117543	GO:0032259	methylation	*DPH5*
ENSG00000138036	GO:0003774	motor activity	*DYNC2LI1*
ENSG00000138036	GO:0005515	protein binding	*DYNC2LI1*
ENSG00000138036	GO:0005829	cytosol	*DYNC2LI1*
ENSG00000138036	GO:0005858	axonemal dynein complex	*DYNC2LI1*
ENSG00000138036	GO:0005881	cytoplasmic microtubule	*DYNC2LI1*
ENSG00000138036	GO:0005930	axoneme	*DYNC2LI1*
ENSG00000138036	GO:0007368	determination of left/right symmetry	*DYNC2LI1*
ENSG00000138036	GO:0008152	metabolic process	*DYNC2LI1*
ENSG00000138036	GO:0019886	antigen processing and presentation of exogenous peptide antigen via MHC class II	*DYNC2LI1*
ENSG00000138036	GO:0030990	intraciliary transport particle	*DYNC2LI1*
ENSG00000138036	GO:0031512	motile primary cilium	*DYNC2LI1*
ENSG00000138036	GO:0036064	ciliary basal body	*DYNC2LI1*
ENSG00000138036	GO:0042384	cilium assembly	*DYNC2LI1*
ENSG00000138036	GO:0045177	apical part of cell	*DYNC2LI1*
ENSG00000138036	GO:0072372	primary cilium	*DYNC2LI1*
ENSG00000186197	GO:0001942	hair follicle development	*EDARADD*
ENSG00000186197	GO:0005515	protein binding	*EDARADD*

ENSG00000186197	GO:0005737	cytoplasm	*EDARADD*
ENSG00000186197	GO:0007165	signal transduction	*EDARADD*
ENSG00000186197	GO:0030154	cell differentiation	*EDARADD*
ENSG00000186197	GO:0042475	odontogenesis of dentin-containing tooth	*EDARADD*
ENSG00000186197	GO:0061153	trachea gland development	*EDARADD*
ENSG00000078401	GO:0000122	negative regulation of transcription from RNA polymerase II promoter	*EDN1*
ENSG00000078401	GO:0001501	skeletal system development	*EDN1*
ENSG00000078401	GO:0001516	prostaglandin biosynthetic process	*EDN1*
ENSG00000078401	GO:0001569	patterning of blood vessels	*EDN1*
ENSG00000078401	GO:0001666	response to hypoxia	*EDN1*
ENSG00000078401	GO:0001701	in utero embryonic development	*EDN1*
ENSG00000078401	GO:0001821	histamine secretion	*EDN1*
ENSG00000078401	GO:0003100	regulation of systemic arterial blood pressure by endothelin	*EDN1*
ENSG00000078401	GO:0005102	receptor binding	*EDN1*
ENSG00000078401	GO:0005125	cytokine activity	*EDN1*
ENSG00000078401	GO:0005179	hormone activity	*EDN1*
ENSG00000078401	GO:0005515	protein binding	*EDN1*
ENSG00000078401	GO:0005576	extracellular region	*EDN1*
ENSG00000078401	GO:0005615	extracellular space	*EDN1*
ENSG00000078401	GO:0005737	cytoplasm	*EDN1*
ENSG00000078401	GO:0006885	regulation of pH	*EDN1*
ENSG00000078401	GO:0007166	cell surface receptor signaling pathway	*EDN1*
ENSG00000078401	GO:0007186	G-protein coupled receptor signaling pathway	*EDN1*
ENSG00000078401	GO:0007204	positive regulation of cytosolic calcium ion concentration	*EDN1*
ENSG00000078401	GO:0007205	protein kinase C-activating G-protein coupled receptor signaling pathway	*EDN1*
ENSG00000078401	GO:0007267	cell-cell signaling	*EDN1*
ENSG00000078401	GO:0007507	heart development	*EDN1*
ENSG00000078401	GO:0007585	respiratory gaseous exchange	*EDN1*
ENSG00000078401	GO:0007589	body fluid secretion	*EDN1*
ENSG00000078401	GO:0008217	regulation of blood pressure	*EDN1*
ENSG00000078401	GO:0008284	positive regulation of cell proliferation	*EDN1*
ENSG00000078401	GO:0009953	dorsal/ventral pattern formation	*EDN1*
ENSG00000078401	GO:0010193	response to ozone	*EDN1*
ENSG00000078401	GO:0010259	multicellular organismal aging	*EDN1*
ENSG00000078401	GO:0010460	positive regulation of heart rate	*EDN1*
ENSG00000078401	GO:0010595	positive regulation of endothelial cell migration	*EDN1*
ENSG00000078401	GO:0010613	positive regulation of cardiac muscle hypertrophy	*EDN1*
ENSG00000078401	GO:0010870	positive regulation of receptor biosynthetic process	*EDN1*
ENSG00000078401	GO:0014032	neural crest cell development	*EDN1*
ENSG00000078401	GO:0014065	phosphatidylinositol 3-kinase signaling	*EDN1*
ENSG00000078401	GO:0014823	response to activity	*EDN1*
ENSG00000078401	GO:0014824	artery smooth muscle contraction	*EDN1*
ENSG00000078401	GO:0014826	vein smooth muscle contraction	*EDN1*
ENSG00000078401	GO:0015758	glucose transport	*EDN1*
ENSG00000078401	GO:0016049	cell growth	*EDN1*
ENSG00000078401	GO:0019229	regulation of vasoconstriction	*EDN1*
ENSG00000078401	GO:0019233	sensory perception of pain	*EDN1*

ENSG00000078401	GO:0019722	calcium-mediated signaling	*EDN1*
ENSG00000078401	GO:0030072	peptide hormone secretion	*EDN1*
ENSG00000078401	GO:0030185	nitric oxide transport	*EDN1*
ENSG00000078401	GO:0030195	negative regulation of blood coagulation	*EDN1*
ENSG00000078401	GO:0030335	positive regulation of cell migration	*EDN1*
ENSG00000078401	GO:0030818	negative regulation of cAMP biosynthetic process	*EDN1*
ENSG00000078401	GO:0031583	phospholipase D-activating G-protein coupled receptor signaling pathway	*EDN1*
ENSG00000078401	GO:0031707	endothelin A receptor binding	*EDN1*
ENSG00000078401	GO:0031708	endothelin B receptor binding	*EDN1*
ENSG00000078401	GO:0032269	negative regulation of cellular protein metabolic process	*EDN1*
ENSG00000078401	GO:0032308	positive regulation of prostaglandin secretion	*EDN1*
ENSG00000078401	GO:0032496	response to lipopolysaccharide	*EDN1*
ENSG00000078401	GO:0033093	Weibel-Palade body	*EDN1*
ENSG00000078401	GO:0033574	response to testosterone	*EDN1*
ENSG00000078401	GO:0034392	negative regulation of smooth muscle cell apoptotic process	*EDN1*
ENSG00000078401	GO:0034696	response to prostaglandin F	*EDN1*
ENSG00000078401	GO:0035094	response to nicotine	*EDN1*
ENSG00000078401	GO:0035556	intracellular signal transduction	*EDN1*
ENSG00000078401	GO:0035690	cellular response to drug	*EDN1*
ENSG00000078401	GO:0035810	positive regulation of urine volume	*EDN1*
ENSG00000078401	GO:0035815	positive regulation of renal sodium excretion	*EDN1*
ENSG00000078401	GO:0035994	response to muscle stretch	*EDN1*
ENSG00000078401	GO:0042045	epithelial fluid transport	*EDN1*
ENSG00000078401	GO:0042310	vasoconstriction	*EDN1*
ENSG00000078401	GO:0042313	protein kinase C deactivation	*EDN1*
ENSG00000078401	GO:0042474	middle ear morphogenesis	*EDN1*
ENSG00000078401	GO:0042482	positive regulation of odontogenesis	*EDN1*
ENSG00000078401	GO:0042493	response to drug	*EDN1*
ENSG00000078401	GO:0042554	superoxide anion generation	*EDN1*
ENSG00000078401	GO:0043179	rhythmic excitation	*EDN1*
ENSG00000078401	GO:0043200	response to amino acid	*EDN1*
ENSG00000078401	GO:0043406	positive regulation of MAP kinase activity	*EDN1*
ENSG00000078401	GO:0043507	positive regulation of JUN kinase activity	*EDN1*
ENSG00000078401	GO:0044321	response to leptin	*EDN1*
ENSG00000078401	GO:0045178	basal part of cell	*EDN1*
ENSG00000078401	GO:0045321	leukocyte activation	*EDN1*
ENSG00000078401	GO:0045429	positive regulation of nitric oxide biosynthetic process	*EDN1*
ENSG00000078401	GO:0045793	positive regulation of cell size	*EDN1*
ENSG00000078401	GO:0045840	positive regulation of mitosis	*EDN1*
ENSG00000078401	GO:0045944	positive regulation of transcription from RNA polymerase II promoter	*EDN1*
ENSG00000078401	GO:0045987	positive regulation of smooth muscle contraction	*EDN1*
ENSG00000078401	GO:0046887	positive regulation of hormone secretion	*EDN1*
ENSG00000078401	GO:0046888	negative regulation of hormone secretion	*EDN1*

ENSG00000078401	GO:0048016	inositol phosphate-mediated signaling	*EDN1*
ENSG00000078401	GO:0048237	rough endoplasmic reticulum lumen	*EDN1*
ENSG00000078401	GO:0048514	blood vessel morphogenesis	*EDN1*
ENSG00000078401	GO:0048661	positive regulation of smooth muscle cell proliferation	*EDN1*
ENSG00000078401	GO:0051216	cartilage development	*EDN1*
ENSG00000078401	GO:0051482	positive regulation of cytosolic calcium ion concentration involved in phospholipase C-activating G-protein coupled signaling pathway	*EDN1*
ENSG00000078401	GO:0051771	negative regulation of nitric-oxide synthase biosynthetic process	*EDN1*
ENSG00000078401	GO:0051899	membrane depolarization	*EDN1*
ENSG00000078401	GO:0051930	regulation of sensory perception of pain	*EDN1*
ENSG00000078401	GO:0060137	maternal process involved in parturition	*EDN1*
ENSG00000078401	GO:0060298	positive regulation of sarcomere organization	*EDN1*
ENSG00000078401	GO:0060585	positive regulation of prostaglandin-endoperoxide synthase activity	*EDN1*
ENSG00000078401	GO:0070101	positive regulation of chemokine-mediated signaling pathway	*EDN1*
ENSG00000078401	GO:0071277	cellular response to calcium ion	*EDN1*
ENSG00000078401	GO:0071346	cellular response to interferon-gamma	*EDN1*
ENSG00000078401	GO:0071347	cellular response to interleukin-1	*EDN1*
ENSG00000078401	GO:0071356	cellular response to tumor necrosis factor	*EDN1*
ENSG00000078401	GO:0071375	cellular response to peptide hormone stimulus	*EDN1*
ENSG00000078401	GO:0071385	cellular response to glucocorticoid stimulus	*EDN1*
ENSG00000078401	GO:0071389	cellular response to mineralocorticoid stimulus	*EDN1*
ENSG00000078401	GO:0071398	cellular response to fatty acid	*EDN1*
ENSG00000078401	GO:0071456	cellular response to hypoxia	*EDN1*
ENSG00000078401	GO:0071548	response to dexamethasone	*EDN1*
ENSG00000078401	GO:0071559	response to transforming growth factor beta	*EDN1*
ENSG00000078401	GO:0071560	cellular response to transforming growth factor beta stimulus	*EDN1*
ENSG00000078401	GO:0090023	positive regulation of neutrophil chemotaxis	*EDN1*
ENSG00000078401	GO:1902074	response to salt	*EDN1*
ENSG00000164307	GO:0001525	angiogenesis	*ERAP1*
ENSG00000164307	GO:0002474	antigen processing and presentation of peptide antigen via MHC class I	*ERAP1*
ENSG00000164307	GO:0004177	aminopeptidase activity	*ERAP1*
ENSG00000164307	GO:0005138	interleukin-6 receptor binding	*ERAP1*
ENSG00000164307	GO:0005151	interleukin-1, Type II receptor binding	*ERAP1*
ENSG00000164307	GO:0005515	protein binding	*ERAP1*
ENSG00000164307	GO:0005576	extracellular region	*ERAP1*
ENSG00000164307	GO:0005615	extracellular space	*ERAP1*
ENSG00000164307	GO:0005737	cytoplasm	*ERAP1*
ENSG00000164307	GO:0005783	endoplasmic reticulum	*ERAP1*
ENSG00000164307	GO:0005788	endoplasmic reticulum lumen	*ERAP1*
ENSG00000164307	GO:0005789	endoplasmic reticulum membrane	*ERAP1*
ENSG00000164307	GO:0005829	cytosol	*ERAP1*

ENSG00000164307	GO:0006508	proteolysis	*ERAP1*
ENSG00000164307	GO:0006509	membrane protein ectodomain proteolysis	*ERAP1*
ENSG00000164307	GO:0008217	regulation of blood pressure	*ERAP1*
ENSG00000164307	GO:0008235	metalloexopeptidase activity	*ERAP1*
ENSG00000164307	GO:0008237	metallopeptidase activity	*ERAP1*
ENSG00000164307	GO:0008270	zinc ion binding	*ERAP1*
ENSG00000164307	GO:0009617	response to bacterium	*ERAP1*
ENSG00000164307	GO:0016020	membrane	*ERAP1*
ENSG00000164307	GO:0016021	integral component of membrane	*ERAP1*
ENSG00000164307	GO:0019885	antigen processing and presentation of endogenous peptide antigen via MHC class I	*ERAP1*
ENSG00000164307	GO:0045088	regulation of innate immune response	*ERAP1*
ENSG00000164307	GO:0045444	fat cell differentiation	*ERAP1*
ENSG00000164307	GO:0045766	positive regulation of angiogenesis	*ERAP1*
ENSG00000164307	GO:0070062	extracellular vesicular exosome	*ERAP1*
ENSG00000082805	GO:0000139	Golgi membrane	*ERC1*
ENSG00000082805	GO:0005515	protein binding	*ERC1*
ENSG00000082805	GO:0005737	cytoplasm	*ERC1*
ENSG00000082805	GO:0006355	regulation of transcription, DNA-templated	*ERC1*
ENSG00000082805	GO:0007252	I-kappaB phosphorylation	*ERC1*
ENSG00000082805	GO:0007275	multicellular organismal development	*ERC1*
ENSG00000082805	GO:0008385	IkappaB kinase complex	*ERC1*
ENSG00000082805	GO:0015031	protein transport	*ERC1*
ENSG00000082805	GO:0017137	Rab GTPase binding	*ERC1*
ENSG00000082805	GO:0042147	retrograde transport, endosome to Golgi	*ERC1*
ENSG00000082805	GO:0042734	presynaptic membrane	*ERC1*
ENSG00000082805	GO:0043522	leucine zipper domain binding	*ERC1*
ENSG00000082805	GO:0045202	synapse	*ERC1*
ENSG00000082805	GO:0051092	positive regulation of NF-kappaB transcription factor activity	*ERC1*
ENSG00000064655	GO:0000287	magnesium ion binding	*EYA2*
ENSG00000064655	GO:0004725	protein tyrosine phosphatase activity	*EYA2*
ENSG00000064655	GO:0005515	protein binding	*EYA2*
ENSG00000064655	GO:0005634	nucleus	*EYA2*
ENSG00000064655	GO:0005737	cytoplasm	*EYA2*
ENSG00000064655	GO:0005813	centrosome	*EYA2*
ENSG00000064655	GO:0006281	DNA repair	*EYA2*
ENSG00000064655	GO:0006351	transcription, DNA-templated	*EYA2*
ENSG00000064655	GO:0006355	regulation of transcription, DNA-templated	*EYA2*
ENSG00000064655	GO:0006470	protein dephosphorylation	*EYA2*
ENSG00000064655	GO:0007275	multicellular organismal development	*EYA2*
ENSG00000064655	GO:0007501	mesodermal cell fate specification	*EYA2*
ENSG00000064655	GO:0014706	striated muscle tissue development	*EYA2*
ENSG00000064655	GO:0016576	histone dephosphorylation	*EYA2*
ENSG00000064655	GO:0035335	peptidyl-tyrosine dephosphorylation	*EYA2*
ENSG00000064655	GO:0097192	extrinsic apoptotic signaling pathway in absence of ligand	*EYA2*
ENSG00000064655	GO:0097345	mitochondrial outer membrane permeabilization	*EYA2*
ENSG00000026103	GO:0002377	immunoglobulin production	*FAS*
ENSG00000026103	GO:0003014	renal system process	*FAS*

ENSG00000026103	GO:0004871	signal transducer activity	*FAS*
ENSG00000026103	GO:0004872	receptor activity	*FAS*
ENSG00000026103	GO:0004888	transmembrane signaling receptor activity	*FAS*
ENSG00000026103	GO:0005515	protein binding	*FAS*
ENSG00000026103	GO:0005576	extracellular region	*FAS*
ENSG00000026103	GO:0005634	nucleus	*FAS*
ENSG00000026103	GO:0005737	cytoplasm	*FAS*
ENSG00000026103	GO:0005829	cytosol	*FAS*
ENSG00000026103	GO:0005886	plasma membrane	*FAS*
ENSG00000026103	GO:0006461	protein complex assembly	*FAS*
ENSG00000026103	GO:0006915	apoptotic process	*FAS*
ENSG00000026103	GO:0006919	activation of cysteine-type endopeptidase activity involved in apoptotic process	*FAS*
ENSG00000026103	GO:0006924	activation-induced cell death of T cells	*FAS*
ENSG00000026103	GO:0006925	inflammatory cell apoptotic process	*FAS*
ENSG00000026103	GO:0006927	transformed cell apoptotic process	*FAS*
ENSG00000026103	GO:0006955	immune response	*FAS*
ENSG00000026103	GO:0007165	signal transduction	*FAS*
ENSG00000026103	GO:0007623	circadian rhythm	*FAS*
ENSG00000026103	GO:0008625	extrinsic apoptotic signaling pathway via death domain receptors	*FAS*
ENSG00000026103	GO:0009636	response to toxic substance	*FAS*
ENSG00000026103	GO:0009897	external side of plasma membrane	*FAS*
ENSG00000026103	GO:0009986	cell surface	*FAS*
ENSG00000026103	GO:0010467	gene expression	*FAS*
ENSG00000026103	GO:0016021	integral component of membrane	*FAS*
ENSG00000026103	GO:0019724	B cell mediated immunity	*FAS*
ENSG00000026103	GO:0019900	kinase binding	*FAS*
ENSG00000026103	GO:0031264	death-inducing signaling complex	*FAS*
ENSG00000026103	GO:0031265	CD95 death-inducing signaling complex	*FAS*
ENSG00000026103	GO:0032464	positive regulation of protein homooligomerization	*FAS*
ENSG00000026103	GO:0042802	identical protein binding	*FAS*
ENSG00000026103	GO:0042981	regulation of apoptotic process	*FAS*
ENSG00000026103	GO:0043029	T cell homeostasis	*FAS*
ENSG00000026103	GO:0043065	positive regulation of apoptotic process	*FAS*
ENSG00000026103	GO:0043066	negative regulation of apoptotic process	*FAS*
ENSG00000026103	GO:0045060	negative thymic T cell selection	*FAS*
ENSG00000026103	GO:0045121	membrane raft	*FAS*
ENSG00000026103	GO:0045619	regulation of lymphocyte differentiation	*FAS*
ENSG00000026103	GO:0045637	regulation of myeloid cell differentiation	*FAS*
ENSG00000026103	GO:0048536	spleen development	*FAS*
ENSG00000026103	GO:0050869	negative regulation of B cell activation	*FAS*
ENSG00000026103	GO:0051260	protein homooligomerization	*FAS*
ENSG00000026103	GO:0051384	response to glucocorticoid	*FAS*
ENSG00000026103	GO:0051402	neuron apoptotic process	*FAS*
ENSG00000026103	GO:0070062	extracellular vesicular exosome	*FAS*
ENSG00000026103	GO:0071260	cellular response to mechanical stimulus	*FAS*
ENSG00000026103	GO:0071285	cellular response to lithium ion	*FAS*

ENSG00000026103	GO:0071455	cellular response to hyperoxia	*FAS*
ENSG00000026103	GO:0097049	motor neuron apoptotic process	*FAS*
ENSG00000026103	GO:0097190	apoptotic signaling pathway	*FAS*
ENSG00000026103	GO:0097191	extrinsic apoptotic signaling pathway	*FAS*
ENSG00000026103	GO:0097192	extrinsic apoptotic signaling pathway in absence of ligand	*FAS*
ENSG00000026103	GO:0097527	necroptotic signaling pathway	*FAS*
ENSG00000026103	GO:2001239	regulation of extrinsic apoptotic signaling pathway in absence of ligand	*FAS*
ENSG00000026103	GO:2001241	positive regulation of extrinsic apoptotic signaling pathway in absence of ligand	*FAS*
ENSG00000183580	GO:0000086	G2/M transition of mitotic cell cycle	*FBXL7*
ENSG00000183580	GO:0000151	ubiquitin ligase complex	*FBXL7*
ENSG00000183580	GO:0004842	ubiquitin-protein transferase activity	*FBXL7*
ENSG00000183580	GO:0005515	protein binding	*FBXL7*
ENSG00000183580	GO:0005737	cytoplasm	*FBXL7*
ENSG00000183580	GO:0005813	centrosome	*FBXL7*
ENSG00000183580	GO:0006511	ubiquitin-dependent protein catabolic process	*FBXL7*
ENSG00000183580	GO:0007067	mitotic nuclear division	*FBXL7*
ENSG00000183580	GO:0008283	cell proliferation	*FBXL7*
ENSG00000183580	GO:0016567	protein ubiquitination	*FBXL7*
ENSG00000183580	GO:0019005	SCF ubiquitin ligase complex	*FBXL7*
ENSG00000183580	GO:0031146	SCF-dependent proteasomal ubiquitin-dependent protein catabolic process	*FBXL7*
ENSG00000189283	GO:0000166	nucleotide binding	*FHIT*
ENSG00000189283	GO:0003824	catalytic activity	*FHIT*
ENSG00000189283	GO:0005515	protein binding	*FHIT*
ENSG00000189283	GO:0005634	nucleus	*FHIT*
ENSG00000189283	GO:0005737	cytoplasm	*FHIT*
ENSG00000189283	GO:0005739	mitochondrion	*FHIT*
ENSG00000189283	GO:0005829	cytosol	*FHIT*
ENSG00000189283	GO:0006163	purine nucleotide metabolic process	*FHIT*
ENSG00000189283	GO:0006260	DNA replication	*FHIT*
ENSG00000189283	GO:0006351	transcription, DNA-templated	*FHIT*
ENSG00000189283	GO:0006355	regulation of transcription, DNA-templated	*FHIT*
ENSG00000189283	GO:0009117	nucleotide metabolic process	*FHIT*
ENSG00000189283	GO:0016787	hydrolase activity	*FHIT*
ENSG00000189283	GO:0031625	ubiquitin protein ligase binding	*FHIT*
ENSG00000189283	GO:0032435	negative regulation of proteasomal ubiquitin-dependent protein catabolic process	*FHIT*
ENSG00000189283	GO:0042802	identical protein binding	*FHIT*
ENSG00000189283	GO:0047710	bis(5'-adenosyl)-triphosphatase activity	*FHIT*
ENSG00000189283	GO:0070062	extracellular vesicular exosome	*FHIT*
ENSG00000189283	GO:0072332	intrinsic apoptotic signaling pathway by p53 class mediator	*FHIT*
ENSG00000185070	GO:0005057	receptor signaling protein activity	*FLRT2*
ENSG00000185070	GO:0005515	protein binding	*FLRT2*
ENSG00000185070	GO:0005578	proteinaceous extracellular matrix	*FLRT2*
ENSG00000185070	GO:0005615	extracellular space	*FLRT2*
ENSG00000185070	GO:0005887	integral component of plasma	*FLRT2*

		membrane	
ENSG00000185070	GO:0007155	cell adhesion	*FLRT2*
ENSG00000185070	GO:0007411	axon guidance	*FLRT2*
ENSG00000185070	GO:0008150	biological_process	*FLRT2*
ENSG00000185070	GO:0030674	protein binding, bridging	*FLRT2*
ENSG00000185070	GO:0035556	intracellular signal transduction	*FLRT2*
ENSG00000185070	GO:0045499	chemorepellent activity	*FLRT2*
ENSG00000185070	GO:0050919	negative chemotaxis	*FLRT2*
ENSG00000185070	GO:0070062	extracellular vesicular exosome	*FLRT2*
ENSG00000185070	GO:2001222	regulation of neuron migration	*FLRT2*
ENSG00000125848	GO:0005057	receptor signaling protein activity	*FLRT3*
ENSG00000125848	GO:0005515	protein binding	*FLRT3*
ENSG00000125848	GO:0005578	proteinaceous extracellular matrix	*FLRT3*
ENSG00000125848	GO:0005615	extracellular space	*FLRT3*
ENSG00000125848	GO:0005887	integral component of plasma membrane	*FLRT3*
ENSG00000125848	GO:0007155	cell adhesion	*FLRT3*
ENSG00000125848	GO:0007411	axon guidance	*FLRT3*
ENSG00000125848	GO:0008150	biological_process	*FLRT3*
ENSG00000125848	GO:0030674	protein binding, bridging	*FLRT3*
ENSG00000125848	GO:0035556	intracellular signal transduction	*FLRT3*
ENSG00000125848	GO:0045499	chemorepellent activity	*FLRT3*
ENSG00000125848	GO:0050919	negative chemotaxis	*FLRT3*
ENSG00000151474	GO:0005737	cytoplasm	*FRMD4A*
ENSG00000151474	GO:0005856	cytoskeleton	*FRMD4A*
ENSG00000151474	GO:0005923	tight junction	*FRMD4A*
ENSG00000151474	GO:0030674	protein binding, bridging	*FRMD4A*
ENSG00000151474	GO:0090162	establishment of epithelial cell polarity	*FRMD4A*
ENSG00000053108	GO:0005509	calcium ion binding	*FSTL4*
ENSG00000053108	GO:0005515	protein binding	*FSTL4*
ENSG00000053108	GO:0005576	extracellular region	*FSTL4*
ENSG00000168843	GO:0005509	calcium ion binding	*FSTL5*
ENSG00000168843	GO:0005515	protein binding	*FSTL5*
ENSG00000168843	GO:0005576	extracellular region	*FSTL5*
ENSG00000082074	GO:0005102	receptor binding	*FYB*
ENSG00000082074	GO:0005515	protein binding	*FYB*
ENSG00000082074	GO:0005634	nucleus	*FYB*
ENSG00000082074	GO:0005829	cytosol	*FYB*
ENSG00000082074	GO:0006468	protein phosphorylation	*FYB*
ENSG00000082074	GO:0006607	NLS-bearing protein import into nucleus	*FYB*
ENSG00000082074	GO:0006955	immune response	*FYB*
ENSG00000082074	GO:0007165	signal transduction	*FYB*
ENSG00000082074	GO:0008150	biological_process	*FYB*
ENSG00000082074	GO:0015629	actin cytoskeleton	*FYB*
ENSG00000082074	GO:0032403	protein complex binding	*FYB*
ENSG00000082074	GO:0035556	intracellular signal transduction	*FYB*
ENSG00000082074	GO:0050852	T cell receptor signaling pathway	*FYB*
ENSG00000094755	GO:0004890	GABA-A receptor activity	*GABRP*
ENSG00000094755	GO:0005230	extracellular ligand-gated ion channel activity	*GABRP*
ENSG00000094755	GO:0005254	chloride channel activity	*GABRP*
ENSG00000094755	GO:0005886	plasma membrane	*GABRP*
ENSG00000094755	GO:0006810	transport	*GABRP*
ENSG00000094755	GO:0006811	ion transport	*GABRP*
ENSG00000094755	GO:0007165	signal transduction	*GABRP*

ENSG00000094755	GO:0016020	membrane	*GABRP*
ENSG00000094755	GO:0016021	integral component of membrane	*GABRP*
ENSG00000094755	GO:0030054	cell junction	*GABRP*
ENSG00000094755	GO:0034220	ion transmembrane transport	*GABRP*
ENSG00000094755	GO:0034707	chloride channel complex	*GABRP*
ENSG00000094755	GO:0045202	synapse	*GABRP*
ENSG00000094755	GO:0045211	postsynaptic membrane	*GABRP*
ENSG00000094755	GO:1902476	chloride transmembrane transport	*GABRP*
ENSG00000168505	GO:0001569	patterning of blood vessels	*GBX2*
ENSG00000168505	GO:0001755	neural crest cell migration	*GBX2*
ENSG00000168505	GO:0003677	DNA binding	*GBX2*
ENSG00000168505	GO:0003700	sequence-specific DNA binding transcription factor activity	*GBX2*
ENSG00000168505	GO:0005634	nucleus	*GBX2*
ENSG00000168505	GO:0006351	transcription, DNA-templated	*GBX2*
ENSG00000168505	GO:0006355	regulation of transcription, DNA-templated	*GBX2*
ENSG00000168505	GO:0007399	nervous system development	*GBX2*
ENSG00000168505	GO:0007411	axon guidance	*GBX2*
ENSG00000168505	GO:0008283	cell proliferation	*GBX2*
ENSG00000168505	GO:0021549	cerebellum development	*GBX2*
ENSG00000168505	GO:0021555	midbrain-hindbrain boundary morphogenesis	*GBX2*
ENSG00000168505	GO:0021568	rhombomere 2 development	*GBX2*
ENSG00000168505	GO:0021794	thalamus development	*GBX2*
ENSG00000168505	GO:0021884	forebrain neuron development	*GBX2*
ENSG00000168505	GO:0021930	cerebellar granule cell precursor proliferation	*GBX2*
ENSG00000168505	GO:0030902	hindbrain development	*GBX2*
ENSG00000168505	GO:0030917	midbrain-hindbrain boundary development	*GBX2*
ENSG00000168505	GO:0035239	tube morphogenesis	*GBX2*
ENSG00000168505	GO:0042472	inner ear morphogenesis	*GBX2*
ENSG00000168505	GO:0043565	sequence-specific DNA binding	*GBX2*
ENSG00000168505	GO:0048483	autonomic nervous system development	*GBX2*
ENSG00000131459	GO:0004360	glutamine-fructose-6-phosphate transaminase (isomerizing) activity	*GFPT2*
ENSG00000131459	GO:0005829	cytosol	*GFPT2*
ENSG00000131459	GO:0005975	carbohydrate metabolic process	*GFPT2*
ENSG00000131459	GO:0006002	fructose 6-phosphate metabolic process	*GFPT2*
ENSG00000131459	GO:0006048	UDP-N-acetylglucosamine biosynthetic process	*GFPT2*
ENSG00000131459	GO:0006112	energy reserve metabolic process	*GFPT2*
ENSG00000131459	GO:0006488	dolichol-linked oligosaccharide biosynthetic process	*GFPT2*
ENSG00000131459	GO:0006541	glutamine metabolic process	*GFPT2*
ENSG00000131459	GO:0008152	metabolic process	*GFPT2*
ENSG00000131459	GO:0016051	carbohydrate biosynthetic process	*GFPT2*
ENSG00000131459	GO:0018279	protein N-linked glycosylation via asparagine	*GFPT2*
ENSG00000131459	GO:0030246	carbohydrate binding	*GFPT2*
ENSG00000131459	GO:0043687	post-translational protein modification	*GFPT2*
ENSG00000131459	GO:0044267	cellular protein metabolic process	*GFPT2*
ENSG00000106571	GO:0000122	negative regulation of transcription from RNA polymerase II promoter	*GLI3*

ENSG00000106571	GO:0000978	RNA polymerase II core promoter proximal region sequence-specific DNA binding	*GLI3*
ENSG00000106571	GO:0001656	metanephros development	*GLI3*
ENSG00000106571	GO:0001658	branching involved in ureteric bud morphogenesis	*GLI3*
ENSG00000106571	GO:0001701	in utero embryonic development	*GLI3*
ENSG00000106571	GO:0001822	kidney development	*GLI3*
ENSG00000106571	GO:0002052	positive regulation of neuroblast proliferation	*GLI3*
ENSG00000106571	GO:0003682	chromatin binding	*GLI3*
ENSG00000106571	GO:0003700	sequence-specific DNA binding transcription factor activity	*GLI3*
ENSG00000106571	GO:0005515	protein binding	*GLI3*
ENSG00000106571	GO:0005634	nucleus	*GLI3*
ENSG00000106571	GO:0005737	cytoplasm	*GLI3*
ENSG00000106571	GO:0005829	cytosol	*GLI3*
ENSG00000106571	GO:0005929	cilium	*GLI3*
ENSG00000106571	GO:0006351	transcription, DNA-templated	*GLI3*
ENSG00000106571	GO:0006355	regulation of transcription, DNA-templated	*GLI3*
ENSG00000106571	GO:0007224	smoothened signaling pathway	*GLI3*
ENSG00000106571	GO:0007389	pattern specification process	*GLI3*
ENSG00000106571	GO:0007411	axon guidance	*GLI3*
ENSG00000106571	GO:0007417	central nervous system development	*GLI3*
ENSG00000106571	GO:0007420	brain development	*GLI3*
ENSG00000106571	GO:0007442	hindgut morphogenesis	*GLI3*
ENSG00000106571	GO:0007507	heart development	*GLI3*
ENSG00000106571	GO:0008013	beta-catenin binding	*GLI3*
ENSG00000106571	GO:0008285	negative regulation of cell proliferation	*GLI3*
ENSG00000106571	GO:0009952	anterior/posterior pattern specification	*GLI3*
ENSG00000106571	GO:0009953	dorsal/ventral pattern formation	*GLI3*
ENSG00000106571	GO:0009954	proximal/distal pattern formation	*GLI3*
ENSG00000106571	GO:0010468	regulation of gene expression	*GLI3*
ENSG00000106571	GO:0016485	protein processing	*GLI3*
ENSG00000106571	GO:0016592	mediator complex	*GLI3*
ENSG00000106571	GO:0016607	nuclear speck	*GLI3*
ENSG00000106571	GO:0017053	transcriptional repressor complex	*GLI3*
ENSG00000106571	GO:0021513	spinal cord dorsal/ventral patterning	*GLI3*
ENSG00000106571	GO:0021522	spinal cord motor neuron differentiation	*GLI3*
ENSG00000106571	GO:0021537	telencephalon development	*GLI3*
ENSG00000106571	GO:0021543	pallium development	*GLI3*
ENSG00000106571	GO:0021544	subpallium development	*GLI3*
ENSG00000106571	GO:0021631	optic nerve morphogenesis	*GLI3*
ENSG00000106571	GO:0021766	hippocampus development	*GLI3*
ENSG00000106571	GO:0021775	smoothened signaling pathway involved in ventral spinal cord interneuron specification	*GLI3*
ENSG00000106571	GO:0021776	smoothened signaling pathway involved in spinal cord motor neuron cell fate specification	*GLI3*
ENSG00000106571	GO:0021798	forebrain dorsal/ventral pattern formation	*GLI3*
ENSG00000106571	GO:0021801	cerebral cortex radial glia guided migration	*GLI3*
ENSG00000106571	GO:0021861	forebrain radial glial cell differentiation	*GLI3*

ENSG00000106571	GO:0021915	neural tube development	*GLI3*
ENSG00000106571	GO:0022018	lateral ganglionic eminence cell proliferation	*GLI3*
ENSG00000106571	GO:0030318	melanocyte differentiation	*GLI3*
ENSG00000106571	GO:0030324	lung development	*GLI3*
ENSG00000106571	GO:0030326	embryonic limb morphogenesis	*GLI3*
ENSG00000106571	GO:0030879	mammary gland development	*GLI3*
ENSG00000106571	GO:0030900	forebrain development	*GLI3*
ENSG00000106571	GO:0032332	positive regulation of chondrocyte differentiation	*GLI3*
ENSG00000106571	GO:0033077	T cell differentiation in thymus	*GLI3*
ENSG00000106571	GO:0035035	histone acetyltransferase binding	*GLI3*
ENSG00000106571	GO:0035108	limb morphogenesis	*GLI3*
ENSG00000106571	GO:0035295	tube development	*GLI3*
ENSG00000106571	GO:0042060	wound healing	*GLI3*
ENSG00000106571	GO:0042127	regulation of cell proliferation	*GLI3*
ENSG00000106571	GO:0042307	positive regulation of protein import into nucleus	*GLI3*
ENSG00000106571	GO:0042475	odontogenesis of dentin-containing tooth	*GLI3*
ENSG00000106571	GO:0042733	embryonic digit morphogenesis	*GLI3*
ENSG00000106571	GO:0042826	histone deacetylase binding	*GLI3*
ENSG00000106571	GO:0042981	regulation of apoptotic process	*GLI3*
ENSG00000106571	GO:0043010	camera-type eye development	*GLI3*
ENSG00000106571	GO:0043066	negative regulation of apoptotic process	*GLI3*
ENSG00000106571	GO:0043565	sequence-specific DNA binding	*GLI3*
ENSG00000106571	GO:0043585	nose morphogenesis	*GLI3*
ENSG00000106571	GO:0043586	tongue development	*GLI3*
ENSG00000106571	GO:0043627	response to estrogen	*GLI3*
ENSG00000106571	GO:0045060	negative thymic T cell selection	*GLI3*
ENSG00000106571	GO:0045595	regulation of cell differentiation	*GLI3*
ENSG00000106571	GO:0045596	negative regulation of cell differentiation	*GLI3*
ENSG00000106571	GO:0045665	negative regulation of neuron differentiation	*GLI3*
ENSG00000106571	GO:0045669	positive regulation of osteoblast differentiation	*GLI3*
ENSG00000106571	GO:0045879	negative regulation of smoothened signaling pathway	*GLI3*
ENSG00000106571	GO:0045892	negative regulation of transcription, DNA-templated	*GLI3*
ENSG00000106571	GO:0045893	positive regulation of transcription, DNA-templated	*GLI3*
ENSG00000106571	GO:0045944	positive regulation of transcription from RNA polymerase II promoter	*GLI3*
ENSG00000106571	GO:0046638	positive regulation of alpha-beta T cell differentiation	*GLI3*
ENSG00000106571	GO:0046639	negative regulation of alpha-beta T cell differentiation	*GLI3*
ENSG00000106571	GO:0046872	metal ion binding	*GLI3*
ENSG00000106571	GO:0048557	embryonic digestive tract morphogenesis	*GLI3*
ENSG00000106571	GO:0048566	embryonic digestive tract development	*GLI3*
ENSG00000106571	GO:0048589	developmental growth	*GLI3*
ENSG00000106571	GO:0048593	camera-type eye morphogenesis	*GLI3*
ENSG00000106571	GO:0048598	embryonic morphogenesis	*GLI3*

ENSG00000106571	GO:0048646	anatomical structure formation involved in morphogenesis	*GLI3*
ENSG00000106571	GO:0048663	neuron fate commitment	*GLI3*
ENSG00000106571	GO:0048704	embryonic skeletal system morphogenesis	*GLI3*
ENSG00000106571	GO:0048709	oligodendrocyte differentiation	*GLI3*
ENSG00000106571	GO:0048754	branching morphogenesis of an epithelial tube	*GLI3*
ENSG00000106571	GO:0048839	inner ear development	*GLI3*
ENSG00000106571	GO:0048856	anatomical structure development	*GLI3*
ENSG00000106571	GO:0060021	palate development	*GLI3*
ENSG00000106571	GO:0060173	limb development	*GLI3*
ENSG00000106571	GO:0060364	frontal suture morphogenesis	*GLI3*
ENSG00000106571	GO:0060366	lambdoid suture morphogenesis	*GLI3*
ENSG00000106571	GO:0060367	sagittal suture morphogenesis	*GLI3*
ENSG00000106571	GO:0060594	mammary gland specification	*GLI3*
ENSG00000106571	GO:0060831	smoothened signaling pathway involved in dorsal/ventral neural tube patterning	*GLI3*
ENSG00000106571	GO:0060840	artery development	*GLI3*
ENSG00000106571	GO:0060873	anterior semicircular canal development	*GLI3*
ENSG00000106571	GO:0060875	lateral semicircular canal development	*GLI3*
ENSG00000106571	GO:0061005	cell differentiation involved in kidney development	*GLI3*
ENSG00000106571	GO:0070242	thymocyte apoptotic process	*GLI3*
ENSG00000106571	GO:0072372	primary cilium	*GLI3*
ENSG00000106571	GO:0090090	negative regulation of canonical Wnt signaling pathway	*GLI3*
ENSG00000182512	GO:0005634	nucleus	*GLRX5*
ENSG00000182512	GO:0005739	mitochondrion	*GLRX5*
ENSG00000182512	GO:0009055	electron carrier activity	*GLRX5*
ENSG00000182512	GO:0015035	protein disulfide oxidoreductase activity	*GLRX5*
ENSG00000182512	GO:0030097	hemopoiesis	*GLRX5*
ENSG00000182512	GO:0045454	cell redox homeostasis	*GLRX5*
ENSG00000182512	GO:0046872	metal ion binding	*GLRX5*
ENSG00000182512	GO:0051537	2 iron, 2 sulfur cluster binding	*GLRX5*
ENSG00000182512	GO:0055114	oxidation-reduction process	*GLRX5*
ENSG00000214415	GO:0001580	detection of chemical stimulus involved in sensory perception of bitter taste	*GNAT3*
ENSG00000214415	GO:0001664	G-protein coupled receptor binding	*GNAT3*
ENSG00000214415	GO:0001669	acrosomal vesicle	*GNAT3*
ENSG00000214415	GO:0001750	photoreceptor outer segment	*GNAT3*
ENSG00000214415	GO:0001917	photoreceptor inner segment	*GNAT3*
ENSG00000214415	GO:0003924	GTPase activity	*GNAT3*
ENSG00000214415	GO:0004871	signal transducer activity	*GNAT3*
ENSG00000214415	GO:0005525	GTP binding	*GNAT3*
ENSG00000214415	GO:0005737	cytoplasm	*GNAT3*
ENSG00000214415	GO:0005834	heterotrimeric G-protein complex	*GNAT3*
ENSG00000214415	GO:0005886	plasma membrane	*GNAT3*
ENSG00000214415	GO:0005930	axoneme	*GNAT3*
ENSG00000214415	GO:0006184	GTP catabolic process	*GNAT3*
ENSG00000214415	GO:0007165	signal transduction	*GNAT3*
ENSG00000214415	GO:0007186	G-protein coupled receptor signaling pathway	*GNAT3*

ENSG00000214415	GO:0007188	adenylate cyclase-modulating G-protein coupled receptor signaling pathway	*GNAT3*
ENSG00000214415	GO:0007193	adenylate cyclase-inhibiting G-protein coupled receptor signaling pathway	*GNAT3*
ENSG00000214415	GO:0007268	synaptic transmission	*GNAT3*
ENSG00000214415	GO:0007596	blood coagulation	*GNAT3*
ENSG00000214415	GO:0008020	G-protein coupled photoreceptor activity	*GNAT3*
ENSG00000214415	GO:0009584	detection of visible light	*GNAT3*
ENSG00000214415	GO:0016324	apical plasma membrane	*GNAT3*
ENSG00000214415	GO:0019001	guanyl nucleotide binding	*GNAT3*
ENSG00000214415	GO:0030168	platelet activation	*GNAT3*
ENSG00000214415	GO:0031683	G-protein beta/gamma-subunit complex binding	*GNAT3*
ENSG00000214415	GO:0035094	response to nicotine	*GNAT3*
ENSG00000214415	GO:0043234	protein complex	*GNAT3*
ENSG00000214415	GO:0046872	metal ion binding	*GNAT3*
ENSG00000214415	GO:0050909	sensory perception of taste	*GNAT3*
ENSG00000214415	GO:0050913	sensory perception of bitter taste	*GNAT3*
ENSG00000214415	GO:0050916	sensory perception of sweet taste	*GNAT3*
ENSG00000214415	GO:0050917	sensory perception of umami taste	*GNAT3*
ENSG00000180269	GO:0004930	G-protein coupled receptor activity	*GPR139*
ENSG00000180269	GO:0005886	plasma membrane	*GPR139*
ENSG00000180269	GO:0007186	G-protein coupled receptor signaling pathway	*GPR139*
ENSG00000180269	GO:0007200	phospholipase C-activating G-protein coupled receptor signaling pathway	*GPR139*
ENSG00000180269	GO:0007218	neuropeptide signaling pathway	*GPR139*
ENSG00000180269	GO:0008188	neuropeptide receptor activity	*GPR139*
ENSG00000180269	GO:0016021	integral component of membrane	*GPR139*
ENSG00000180269	GO:0046983	protein dimerization activity	*GPR139*
ENSG00000085382	GO:0000139	Golgi membrane	*HACE1*
ENSG00000085382	GO:0004842	ubiquitin-protein transferase activity	*HACE1*
ENSG00000085382	GO:0005515	protein binding	*HACE1*
ENSG00000085382	GO:0005634	nucleus	*HACE1*
ENSG00000085382	GO:0005783	endoplasmic reticulum	*HACE1*
ENSG00000085382	GO:0006351	transcription, DNA-templated	*HACE1*
ENSG00000085382	GO:0006355	regulation of transcription, DNA-templated	*HACE1*
ENSG00000085382	GO:0007030	Golgi organization	*HACE1*
ENSG00000085382	GO:0007049	cell cycle	*HACE1*
ENSG00000085382	GO:0016567	protein ubiquitination	*HACE1*
ENSG00000085382	GO:0016601	Rac protein signal transduction	*HACE1*
ENSG00000085382	GO:0016874	ligase activity	*HACE1*
ENSG00000085382	GO:0017137	Rab GTPase binding	*HACE1*
ENSG00000085382	GO:0030334	regulation of cell migration	*HACE1*
ENSG00000085382	GO:0032580	Golgi cisterna membrane	*HACE1*
ENSG00000085382	GO:0042787	protein ubiquitination involved in ubiquitin-dependent protein catabolic process	*HACE1*
ENSG00000085382	GO:0048365	Rac GTPase binding	*HACE1*
ENSG00000085382	GO:0061025	membrane fusion	*HACE1*
ENSG00000085382	GO:0070936	protein K48-linked ubiquitination	*HACE1*
ENSG00000232962	GO:0002587	negative regulation of antigen processing and presentation of peptide antigen via MHC class II	*HLA-DOA*

ENSG00000232962	GO:0005515	protein binding	*HLA-DOA*
ENSG00000232962	GO:0005765	lysosomal membrane	*HLA-DOA*
ENSG00000232962	GO:0005886	plasma membrane	*HLA-DOA*
ENSG00000232962	GO:0006955	immune response	*HLA-DOA*
ENSG00000232962	GO:0007165	signal transduction	*HLA-DOA*
ENSG00000232962	GO:0010008	endosome membrane	*HLA-DOA*
ENSG00000232962	GO:0016020	membrane	*HLA-DOA*
ENSG00000232962	GO:0016021	integral component of membrane	*HLA-DOA*
ENSG00000232962	GO:0019882	antigen processing and presentation	*HLA-DOA*
ENSG00000232962	GO:0019886	antigen processing and presentation of exogenous peptide antigen via MHC class II	*HLA-DOA*
ENSG00000232962	GO:0023026	MHC class II protein complex binding	*HLA-DOA*
ENSG00000232962	GO:0032395	MHC class II receptor activity	*HLA-DOA*
ENSG00000232962	GO:0042613	MHC class II protein complex	*HLA-DOA*
ENSG00000232962	GO:0045580	regulation of T cell differentiation	*HLA-DOA*
ENSG00000230763	GO:0000139	Golgi membrane	*HLA-DPB1*
ENSG00000237710	GO:0000139	Golgi membrane	*HLA-DPB1*
ENSG00000226826	GO:0000139	Golgi membrane	*HLA-DPB1*
ENSG00000223865	GO:0000139	Golgi membrane	*HLA-DPB1*
ENSG00000215048	GO:0000139	Golgi membrane	*HLA-DPB1*
ENSG00000236693	GO:0000139	Golgi membrane	*HLA-DPB1*
ENSG00000237710	GO:0002504	antigen processing and presentation of peptide or polysaccharide antigen via MHC class II	*HLA-DPB1*
ENSG00000230763	GO:0005515	protein binding	*HLA-DPB1*
ENSG00000230763	GO:0005765	lysosomal membrane	*HLA-DPB1*
ENSG00000230763	GO:0005886	plasma membrane	*HLA-DPB1*
ENSG00000230763	GO:0006955	immune response	*HLA-DPB1*
ENSG00000230763	GO:0009986	cell surface	*HLA-DPB1*
ENSG00000230763	GO:0010008	endosome membrane	*HLA-DPB1*
ENSG00000230763	GO:0012507	ER to Golgi transport vesicle membrane	*HLA-DPB1*
ENSG00000230763	GO:0016020	membrane	*HLA-DPB1*
ENSG00000237710	GO:0016021	integral component of membrane	*HLA-DPB1*
ENSG00000230763	GO:0019221	cytokine-mediated signaling pathway	*HLA-DPB1*
ENSG00000230763	GO:0019882	antigen processing and presentation	*HLA-DPB1*
ENSG00000230763	GO:0019886	antigen processing and presentation of exogenous peptide antigen via MHC class II	*HLA-DPB1*
ENSG00000230763	GO:0030658	transport vesicle membrane	*HLA-DPB1*
ENSG00000230763	GO:0030666	endocytic vesicle membrane	*HLA-DPB1*
ENSG00000230763	GO:0030669	clathrin-coated endocytic vesicle membrane	*HLA-DPB1*
ENSG00000230763	GO:0031295	T cell costimulation	*HLA-DPB1*
ENSG00000230763	GO:0032588	trans-Golgi network membrane	*HLA-DPB1*
ENSG00000230763	GO:0032729	positive regulation of interferon-gamma production	*HLA-DPB1*
ENSG00000230763	GO:0042102	positive regulation of T cell proliferation	*HLA-DPB1*
ENSG00000230763	GO:0042605	peptide antigen binding	*HLA-DPB1*
ENSG00000230763	GO:0042613	MHC class II protein complex	*HLA-DPB1*
ENSG00000230763	GO:0050852	T cell receptor signaling pathway	*HLA-DPB1*
ENSG00000230763	GO:0050870	positive regulation of T cell activation	*HLA-DPB1*
ENSG00000230763	GO:0060333	interferon-gamma-mediated signaling pathway	*HLA-DPB1*
ENSG00000230763	GO:0071556	integral component of lumenal side of	*HLA-DPB1*

		endoplasmic reticulum membrane	
ENSG00000173083	GO:0004566	beta-glucuronidase activity	*HPSE*
ENSG00000173083	GO:0005515	protein binding	*HPSE*
ENSG00000173083	GO:0005576	extracellular region	*HPSE*
ENSG00000173083	GO:0005634	nucleus	*HPSE*
ENSG00000173083	GO:0005764	lysosome	*HPSE*
ENSG00000173083	GO:0005765	lysosomal membrane	*HPSE*
ENSG00000173083	GO:0005975	carbohydrate metabolic process	*HPSE*
ENSG00000173083	GO:0006027	glycosaminoglycan catabolic process	*HPSE*
ENSG00000173083	GO:0006029	proteoglycan metabolic process	*HPSE*
ENSG00000173083	GO:0007160	cell-matrix adhesion	*HPSE*
ENSG00000173083	GO:0008152	metabolic process	*HPSE*
ENSG00000173083	GO:0010575	positive regulation vascular endothelial growth factor production	*HPSE*
ENSG00000173083	GO:0016020	membrane	*HPSE*
ENSG00000173083	GO:0016798	hydrolase activity, acting on glycosyl bonds	*HPSE*
ENSG00000173083	GO:0030194	positive regulation of blood coagulation	*HPSE*
ENSG00000173083	GO:0030200	heparan sulfate proteoglycan catabolic process	*HPSE*
ENSG00000173083	GO:0030203	glycosaminoglycan metabolic process	*HPSE*
ENSG00000173083	GO:0030305	heparanase activity	*HPSE*
ENSG00000173083	GO:0033690	positive regulation of osteoblast proliferation	*HPSE*
ENSG00000173083	GO:0042060	wound healing	*HPSE*
ENSG00000173083	GO:0043202	lysosomal lumen	*HPSE*
ENSG00000173083	GO:0043231	intracellular membrane-bounded organelle	*HPSE*
ENSG00000173083	GO:0044281	small molecule metabolic process	*HPSE*
ENSG00000173083	GO:0045121	membrane raft	*HPSE*
ENSG00000173083	GO:0045545	syndecan binding	*HPSE*
ENSG00000173083	GO:0046983	protein dimerization activity	*HPSE*
ENSG00000173083	GO:0051797	regulation of hair follicle development	*HPSE*
ENSG00000173083	GO:0051798	positive regulation of hair follicle development	*HPSE*
ENSG00000173083	GO:0051897	positive regulation of protein kinase B signaling	*HPSE*
ENSG00000173083	GO:0060055	angiogenesis involved in wound healing	*HPSE*
ENSG00000173083	GO:0061042	vascular wound healing	*HPSE*
ENSG00000170606	GO:0000902	cell morphogenesis	*HSPA4*
ENSG00000170606	GO:0005515	protein binding	*HSPA4*
ENSG00000170606	GO:0005524	ATP binding	*HSPA4*
ENSG00000170606	GO:0005739	mitochondrion	*HSPA4*
ENSG00000170606	GO:0005829	cytosol	*HSPA4*
ENSG00000170606	GO:0006986	response to unfolded protein	*HSPA4*
ENSG00000170606	GO:0045040	protein import into mitochondrial outer membrane	*HSPA4*
ENSG00000170606	GO:0051131	chaperone-mediated protein complex assembly	*HSPA4*
ENSG00000170606	GO:0070062	extracellular vesicular exosome	*HSPA4*
ENSG00000137959	GO:0005737	cytoplasm	*IFI44L*
ENSG00000137959	GO:0006955	immune response	*IFI44L*
ENSG00000137959	GO:0051607	defense response to virus	*IFI44L*
ENSG00000188487	GO:0000132	establishment of mitotic spindle orientation	*INSC*

ENSG00000188487	GO:0005488	binding	*INSC*
ENSG00000188487	GO:0005515	protein binding	*INSC*
ENSG00000188487	GO:0005737	cytoplasm	*INSC*
ENSG00000188487	GO:0007399	nervous system development	*INSC*
ENSG00000188487	GO:0045177	apical part of cell	*INSC*
ENSG00000188487	GO:0060487	lung epithelial cell differentiation	*INSC*
ENSG00000170561	GO:0001656	metanephros development	*IRX2*
ENSG00000170561	GO:0003677	DNA binding	*IRX2*
ENSG00000170561	GO:0003700	sequence-specific DNA binding transcription factor activity	*IRX2*
ENSG00000170561	GO:0005634	nucleus	*IRX2*
ENSG00000170561	GO:0006355	regulation of transcription, DNA-templated	*IRX2*
ENSG00000170561	GO:0043565	sequence-specific DNA binding	*IRX2*
ENSG00000170561	GO:0072086	specification of loop of Henle identity	*IRX2*
ENSG00000170561	GO:0072272	proximal/distal pattern formation involved in metanephric nephron development	*IRX2*
ENSG00000140015	GO:0000155	phosphorelay sensor kinase activity	*KCNH5*
ENSG00000140015	GO:0000160	phosphorelay signal transduction system	*KCNH5*
ENSG00000140015	GO:0004871	signal transducer activity	*KCNH5*
ENSG00000140015	GO:0005216	ion channel activity	*KCNH5*
ENSG00000140015	GO:0005249	voltage-gated potassium channel activity	*KCNH5*
ENSG00000140015	GO:0005516	calmodulin binding	*KCNH5*
ENSG00000140015	GO:0005622	intracellular	*KCNH5*
ENSG00000140015	GO:0005886	plasma membrane	*KCNH5*
ENSG00000140015	GO:0005887	integral component of plasma membrane	*KCNH5*
ENSG00000140015	GO:0006355	regulation of transcription, DNA-templated	*KCNH5*
ENSG00000140015	GO:0006811	ion transport	*KCNH5*
ENSG00000140015	GO:0007165	signal transduction	*KCNH5*
ENSG00000140015	GO:0007268	synaptic transmission	*KCNH5*
ENSG00000140015	GO:0009986	cell surface	*KCNH5*
ENSG00000140015	GO:0010389	regulation of G2/M transition of mitotic cell cycle	*KCNH5*
ENSG00000140015	GO:0016020	membrane	*KCNH5*
ENSG00000140015	GO:0023014	signal transduction by phosphorylation	*KCNH5*
ENSG00000140015	GO:0034765	regulation of ion transmembrane transport	*KCNH5*
ENSG00000140015	GO:0042391	regulation of membrane potential	*KCNH5*
ENSG00000140015	GO:0055085	transmembrane transport	*KCNH5*
ENSG00000140015	GO:0071805	potassium ion transmembrane transport	*KCNH5*
ENSG00000182132	GO:0005244	voltage-gated ion channel activity	*KCNIP1*
ENSG00000182132	GO:0005267	potassium channel activity	*KCNIP1*
ENSG00000182132	GO:0005509	calcium ion binding	*KCNIP1*
ENSG00000182132	GO:0005513	detection of calcium ion	*KCNIP1*
ENSG00000182132	GO:0005515	protein binding	*KCNIP1*
ENSG00000182132	GO:0005578	proteinaceous extracellular matrix	*KCNIP1*
ENSG00000182132	GO:0005886	plasma membrane	*KCNIP1*
ENSG00000182132	GO:0007165	signal transduction	*KCNIP1*
ENSG00000182132	GO:0007268	synaptic transmission	*KCNIP1*
ENSG00000182132	GO:0015459	potassium channel regulator activity	*KCNIP1*
ENSG00000182132	GO:0030425	dendrite	*KCNIP1*

ENSG00000182132	GO:0034705	potassium channel complex	*KCNIP1*
ENSG00000182132	GO:0034765	regulation of ion transmembrane transport	*KCNIP1*
ENSG00000182132	GO:0043025	neuronal cell body	*KCNIP1*
ENSG00000182132	GO:0044325	ion channel binding	*KCNIP1*
ENSG00000182132	GO:0045760	positive regulation of action potential	*KCNIP1*
ENSG00000182132	GO:0046982	protein heterodimerization activity	*KCNIP1*
ENSG00000182132	GO:0047485	protein N-terminus binding	*KCNIP1*
ENSG00000182132	GO:0071805	potassium ion transmembrane transport	*KCNIP1*
ENSG00000109787	GO:0003677	DNA binding	*KLF3*
ENSG00000109787	GO:0003700	sequence-specific DNA binding transcription factor activity	*KLF3*
ENSG00000109787	GO:0005634	nucleus	*KLF3*
ENSG00000109787	GO:0006351	transcription, DNA-templated	*KLF3*
ENSG00000109787	GO:0006355	regulation of transcription, DNA-templated	*KLF3*
ENSG00000109787	GO:0007275	multicellular organismal development	*KLF3*
ENSG00000109787	GO:0046872	metal ion binding	*KLF3*
ENSG00000109787	GO:1901653	cellular response to peptide	*KLF3*
ENSG00000168216	GO:0005158	insulin receptor binding	*LMBRD1*
ENSG00000168216	GO:0005764	lysosome	*LMBRD1*
ENSG00000168216	GO:0005765	lysosomal membrane	*LMBRD1*
ENSG00000168216	GO:0005886	plasma membrane	*LMBRD1*
ENSG00000168216	GO:0006766	vitamin metabolic process	*LMBRD1*
ENSG00000168216	GO:0006767	water-soluble vitamin metabolic process	*LMBRD1*
ENSG00000168216	GO:0006810	transport	*LMBRD1*
ENSG00000168216	GO:0009235	cobalamin metabolic process	*LMBRD1*
ENSG00000168216	GO:0016020	membrane	*LMBRD1*
ENSG00000168216	GO:0016021	integral component of membrane	*LMBRD1*
ENSG00000168216	GO:0016032	viral process	*LMBRD1*
ENSG00000168216	GO:0031419	cobalamin binding	*LMBRD1*
ENSG00000168216	GO:0038016	insulin receptor internalization	*LMBRD1*
ENSG00000168216	GO:0044281	small molecule metabolic process	*LMBRD1*
ENSG00000168216	GO:0045334	clathrin-coated endocytic vesicle	*LMBRD1*
ENSG00000168216	GO:0046325	negative regulation of glucose import	*LMBRD1*
ENSG00000168216	GO:0046627	negative regulation of insulin receptor signaling pathway	*LMBRD1*
ENSG00000168216	GO:0051898	negative regulation of protein kinase B signaling	*LMBRD1*
ENSG00000081479	GO:0001523	retinoid metabolic process	*LRP2*
ENSG00000081479	GO:0005509	calcium ion binding	*LRP2*
ENSG00000081479	GO:0005515	protein binding	*LRP2*
ENSG00000081479	GO:0005764	lysosome	*LRP2*
ENSG00000081479	GO:0005765	lysosomal membrane	*LRP2*
ENSG00000081479	GO:0005768	endosome	*LRP2*
ENSG00000081479	GO:0005783	endoplasmic reticulum	*LRP2*
ENSG00000081479	GO:0005794	Golgi apparatus	*LRP2*
ENSG00000081479	GO:0005886	plasma membrane	*LRP2*
ENSG00000081479	GO:0005903	brush border	*LRP2*
ENSG00000081479	GO:0005905	coated pit	*LRP2*
ENSG00000081479	GO:0006486	protein glycosylation	*LRP2*
ENSG00000081479	GO:0006629	lipid metabolic process	*LRP2*
ENSG00000081479	GO:0006766	vitamin metabolic process	*LRP2*
ENSG00000081479	GO:0006897	endocytosis	*LRP2*

ENSG00000081479	GO:0006898	receptor-mediated endocytosis	*LRP2*
ENSG00000081479	GO:0007603	phototransduction, visible light	*LRP2*
ENSG00000081479	GO:0008202	steroid metabolic process	*LRP2*
ENSG00000081479	GO:0008283	cell proliferation	*LRP2*
ENSG00000081479	GO:0016020	membrane	*LRP2*
ENSG00000081479	GO:0016021	integral component of membrane	*LRP2*
ENSG00000081479	GO:0016324	apical plasma membrane	*LRP2*
ENSG00000081479	GO:0017124	SH3 domain binding	*LRP2*
ENSG00000081479	GO:0030139	endocytic vesicle	*LRP2*
ENSG00000081479	GO:0030900	forebrain development	*LRP2*
ENSG00000081479	GO:0031526	brush border membrane	*LRP2*
ENSG00000081479	GO:0042359	vitamin D metabolic process	*LRP2*
ENSG00000081479	GO:0043235	receptor complex	*LRP2*
ENSG00000081479	GO:0044281	small molecule metabolic process	*LRP2*
ENSG00000081479	GO:0045177	apical part of cell	*LRP2*
ENSG00000081479	GO:0070062	extracellular vesicular exosome	*LRP2*
ENSG00000172264	GO:0005634	nucleus	*MACROD2*
ENSG00000172264	GO:0006974	cellular response to DNA damage stimulus	*MACROD2*
ENSG00000172264	GO:0007420	brain development	*MACROD2*
ENSG00000172264	GO:0016798	hydrolase activity, acting on glycosyl bonds	*MACROD2*
ENSG00000172264	GO:0019213	deacetylase activity	*MACROD2*
ENSG00000172264	GO:0042278	purine nucleoside metabolic process	*MACROD2*
ENSG00000172264	GO:0051725	protein de-ADP-ribosylation	*MACROD2*
ENSG00000152127	GO:0000139	Golgi membrane	*MGAT5*
ENSG00000152127	GO:0005794	Golgi apparatus	*MGAT5*
ENSG00000152127	GO:0006487	protein N-linked glycosylation	*MGAT5*
ENSG00000152127	GO:0008375	acetylglucosaminyltransferase activity	*MGAT5*
ENSG00000152127	GO:0016021	integral component of membrane	*MGAT5*
ENSG00000152127	GO:0018279	protein N-linked glycosylation via asparagine	*MGAT5*
ENSG00000152127	GO:0030144	alpha-1,6-mannosylglycoprotein 6-beta-N-acetylglucosaminyltransferase activity	*MGAT5*
ENSG00000152127	GO:0043687	post-translational protein modification	*MGAT5*
ENSG00000152127	GO:0044267	cellular protein metabolic process	*MGAT5*
ENSG00000152127	GO:0070062	extracellular vesicular exosome	*MGAT5*
ENSG00000137675	GO:0004222	metalloendopeptidase activity	*MMP27*
ENSG00000137675	GO:0005509	calcium ion binding	*MMP27*
ENSG00000137675	GO:0005578	proteinaceous extracellular matrix	*MMP27*
ENSG00000137675	GO:0006508	proteolysis	*MMP27*
ENSG00000137675	GO:0008237	metallopeptidase activity	*MMP27*
ENSG00000137675	GO:0008270	zinc ion binding	*MMP27*
ENSG00000137675	GO:0030574	collagen catabolic process	*MMP27*
ENSG00000137675	GO:0031012	extracellular matrix	*MMP27*
ENSG00000135698	GO:0000087	mitotic M phase	*MPHOSPH6*
ENSG00000135698	GO:0000176	nuclear exosome (RNase complex)	*MPHOSPH6*
ENSG00000135698	GO:0000178	exosome (RNase complex)	*MPHOSPH6*
ENSG00000135698	GO:0000460	maturation of 5.8S rRNA	*MPHOSPH6*
ENSG00000135698	GO:0003723	RNA binding	*MPHOSPH6*
ENSG00000135698	GO:0005515	protein binding	*MPHOSPH6*
ENSG00000135698	GO:0005634	nucleus	*MPHOSPH6*
ENSG00000135698	GO:0005730	nucleolus	*MPHOSPH6*
ENSG00000135698	GO:0005737	cytoplasm	*MPHOSPH6*
ENSG00000170681	GO:0005737	cytoplasm	*MURC*

ENSG00000170681	GO:0006351	transcription, DNA-templated	*MURC*
ENSG00000170681	GO:0007517	muscle organ development	*MURC*
ENSG00000170681	GO:0030018	Z disc	*MURC*
ENSG00000170681	GO:0030154	cell differentiation	*MURC*
ENSG00000170681	GO:0045944	positive regulation of transcription from RNA polymerase II promoter	*MURC*
ENSG00000034971	GO:0001649	osteoblast differentiation	*MYOC*
ENSG00000034971	GO:0001953	negative regulation of cell-matrix adhesion	*MYOC*
ENSG00000034971	GO:0001968	fibronectin binding	*MYOC*
ENSG00000034971	GO:0005109	frizzled binding	*MYOC*
ENSG00000034971	GO:0005515	protein binding	*MYOC*
ENSG00000034971	GO:0005578	proteinaceous extracellular matrix	*MYOC*
ENSG00000034971	GO:0005615	extracellular space	*MYOC*
ENSG00000034971	GO:0005741	mitochondrial outer membrane	*MYOC*
ENSG00000034971	GO:0005743	mitochondrial inner membrane	*MYOC*
ENSG00000034971	GO:0005758	mitochondrial intermembrane space	*MYOC*
ENSG00000034971	GO:0005783	endoplasmic reticulum	*MYOC*
ENSG00000034971	GO:0005791	rough endoplasmic reticulum	*MYOC*
ENSG00000034971	GO:0005794	Golgi apparatus	*MYOC*
ENSG00000034971	GO:0005929	cilium	*MYOC*
ENSG00000034971	GO:0014068	positive regulation of phosphatidylinositol 3-kinase signaling	*MYOC*
ENSG00000034971	GO:0014734	skeletal muscle hypertrophy	*MYOC*
ENSG00000034971	GO:0016023	cytoplasmic membrane-bounded vesicle	*MYOC*
ENSG00000034971	GO:0022011	myelination in peripheral nervous system	*MYOC*
ENSG00000034971	GO:0030335	positive regulation of cell migration	*MYOC*
ENSG00000034971	GO:0030971	receptor tyrosine kinase binding	*MYOC*
ENSG00000034971	GO:0031012	extracellular matrix	*MYOC*
ENSG00000034971	GO:0031175	neuron projection development	*MYOC*
ENSG00000034971	GO:0031410	cytoplasmic vesicle	*MYOC*
ENSG00000034971	GO:0032027	myosin light chain binding	*MYOC*
ENSG00000034971	GO:0033268	node of Ranvier	*MYOC*
ENSG00000034971	GO:0035024	negative regulation of Rho protein signal transduction	*MYOC*
ENSG00000034971	GO:0038031	non-canonical Wnt signaling pathway via JNK cascade	*MYOC*
ENSG00000034971	GO:0038133	ERBB2-ERBB3 signaling pathway	*MYOC*
ENSG00000034971	GO:0043408	regulation of MAPK cascade	*MYOC*
ENSG00000034971	GO:0045162	clustering of voltage-gated sodium channels	*MYOC*
ENSG00000034971	GO:0051496	positive regulation of stress fiber assembly	*MYOC*
ENSG00000034971	GO:0051497	negative regulation of stress fiber assembly	*MYOC*
ENSG00000034971	GO:0051894	positive regulation of focal adhesion assembly	*MYOC*
ENSG00000034971	GO:0051897	positive regulation of protein kinase B signaling	*MYOC*
ENSG00000034971	GO:0051901	positive regulation of mitochondrial depolarization	*MYOC*
ENSG00000034971	GO:0060348	bone development	*MYOC*
ENSG00000034971	GO:0070062	extracellular vesicular exosome	*MYOC*
ENSG00000034971	GO:1900026	positive regulation of substrate adhesion-dependent cell spreading	*MYOC*

ENSG00000161048	GO:0004620	phospholipase activity	*NAPEPLD*
ENSG00000275723	GO:0008152	metabolic process	*NAPEPLD*
ENSG00000161048	GO:0008152	metabolic process	*NAPEPLD*
ENSG00000275723	GO:0008270	zinc ion binding	*NAPEPLD*
ENSG00000161048	GO:0008270	zinc ion binding	*NAPEPLD*
ENSG00000275723	GO:0009395	phospholipid catabolic process	*NAPEPLD*
ENSG00000161048	GO:0009395	phospholipid catabolic process	*NAPEPLD*
ENSG00000275723	GO:0042622	photoreceptor outer segment membrane	*NAPEPLD*
ENSG00000161048	GO:0042622	photoreceptor outer segment membrane	*NAPEPLD*
ENSG00000275723	GO:0070062	extracellular vesicular exosome	*NAPEPLD*
ENSG00000161048	GO:0070062	extracellular vesicular exosome	*NAPEPLD*
ENSG00000275723	GO:0070290	N-acylphosphatidylethanolamine-specific phospholipase D activity	*NAPEPLD*
ENSG00000161048	GO:0070290	N-acylphosphatidylethanolamine-specific phospholipase D activity	*NAPEPLD*
ENSG00000166833	GO:0003025	regulation of systemic arterial blood pressure by baroreceptor feedback	*NAV2*
ENSG00000166833	GO:0004386	helicase activity	*NAV2*
ENSG00000166833	GO:0005515	protein binding	*NAV2*
ENSG00000166833	GO:0005524	ATP binding	*NAV2*
ENSG00000166833	GO:0005614	interstitial matrix	*NAV2*
ENSG00000166833	GO:0005634	nucleus	*NAV2*
ENSG00000166833	GO:0007399	nervous system development	*NAV2*
ENSG00000166833	GO:0007605	sensory perception of sound	*NAV2*
ENSG00000166833	GO:0007608	sensory perception of smell	*NAV2*
ENSG00000166833	GO:0007626	locomotory behavior	*NAV2*
ENSG00000166833	GO:0008152	metabolic process	*NAV2*
ENSG00000166833	GO:0008201	heparin binding	*NAV2*
ENSG00000166833	GO:0021554	optic nerve development	*NAV2*
ENSG00000166833	GO:0021563	glossopharyngeal nerve development	*NAV2*
ENSG00000166833	GO:0021564	vagus nerve development	*NAV2*
ENSG00000166833	GO:0031012	extracellular matrix	*NAV2*
ENSG00000151779	GO:0000956	nuclear-transcribed mRNA catabolic process	*NBAS*
ENSG00000151779	GO:0005515	protein binding	*NBAS*
ENSG00000151779	GO:0005737	cytoplasm	*NBAS*
ENSG00000151779	GO:0016020	membrane	*NBAS*
ENSG00000151779	GO:2000623	negative regulation of nuclear-transcribed mRNA catabolic process, nonsense-mediated decay	*NBAS*
ENSG00000277586	GO:0000226	microtubule cytoskeleton organization	*NEFL*
ENSG00000277586	GO:0005198	structural molecule activity	*NEFL*
ENSG00000277586	GO:0005200	structural constituent of cytoskeleton	*NEFL*
ENSG00000277586	GO:0005515	protein binding	*NEFL*
ENSG00000277586	GO:0005737	cytoplasm	*NEFL*
ENSG00000277586	GO:0005829	cytosol	*NEFL*
ENSG00000277586	GO:0005882	intermediate filament	*NEFL*
ENSG00000277586	GO:0005883	neurofilament	*NEFL*
ENSG00000277586	GO:0007268	synaptic transmission	*NEFL*
ENSG00000277586	GO:0008022	protein C-terminus binding	*NEFL*
ENSG00000277586	GO:0008089	anterograde axon cargo transport	*NEFL*
ENSG00000277586	GO:0008090	retrograde axon cargo transport	*NEFL*
ENSG00000277586	GO:0008219	cell death	*NEFL*
ENSG00000277586	GO:0009636	response to toxic substance	*NEFL*

ENSG00000277586	GO:0010033	response to organic substance	*NEFL*
ENSG00000277586	GO:0010243	response to organonitrogen compound	*NEFL*
ENSG00000277586	GO:0014012	peripheral nervous system axon regeneration	*NEFL*
ENSG00000277586	GO:0019896	axon transport of mitochondrion	*NEFL*
ENSG00000277586	GO:0019904	protein domain specific binding	*NEFL*
ENSG00000277586	GO:0030424	axon	*NEFL*
ENSG00000277586	GO:0030426	growth cone	*NEFL*
ENSG00000277586	GO:0030674	protein binding, bridging	*NEFL*
ENSG00000277586	GO:0031133	regulation of axon diameter	*NEFL*
ENSG00000277586	GO:0033596	TSC1-TSC2 complex	*NEFL*
ENSG00000277586	GO:0033693	neurofilament bundle assembly	*NEFL*
ENSG00000277586	GO:0040011	locomotion	*NEFL*
ENSG00000277586	GO:0042802	identical protein binding	*NEFL*
ENSG00000277586	GO:0043005	neuron projection	*NEFL*
ENSG00000277586	GO:0043274	phospholipase binding	*NEFL*
ENSG00000277586	GO:0043434	response to peptide hormone	*NEFL*
ENSG00000277586	GO:0043524	negative regulation of neuron apoptotic process	*NEFL*
ENSG00000277586	GO:0045109	intermediate filament organization	*NEFL*
ENSG00000277586	GO:0045110	intermediate filament bundle assembly	*NEFL*
ENSG00000277586	GO:0048812	neuron projection morphogenesis	*NEFL*
ENSG00000277586	GO:0050772	positive regulation of axonogenesis	*NEFL*
ENSG00000277586	GO:0050885	neuromuscular process controlling balance	*NEFL*
ENSG00000277586	GO:0051258	protein polymerization	*NEFL*
ENSG00000277586	GO:0051412	response to corticosterone	*NEFL*
ENSG00000277586	GO:0060052	neurofilament cytoskeleton organization	*NEFL*
ENSG00000147862	GO:0000122	negative regulation of transcription from RNA polymerase II promoter	*NFIB*
ENSG00000147862	GO:0000978	RNA polymerase II core promoter proximal region sequence-specific DNA binding	*NFIB*
ENSG00000147862	GO:0000981	sequence-specific DNA binding RNA polymerase II transcription factor activity	*NFIB*
ENSG00000147862	GO:0001077	RNA polymerase II core promoter proximal region sequence-specific DNA binding transcription factor activity involved in positive regulation of transcription	*NFIB*
ENSG00000147862	GO:0001106	RNA polymerase II transcription corepressor activity	*NFIB*
ENSG00000147862	GO:0002062	chondrocyte differentiation	*NFIB*
ENSG00000147862	GO:0003677	DNA binding	*NFIB*
ENSG00000147862	GO:0003690	double-stranded DNA binding	*NFIB*
ENSG00000147862	GO:0003700	sequence-specific DNA binding transcription factor activity	*NFIB*
ENSG00000147862	GO:0005622	intracellular	*NFIB*
ENSG00000147862	GO:0005634	nucleus	*NFIB*
ENSG00000147862	GO:0005730	nucleolus	*NFIB*
ENSG00000147862	GO:0006260	DNA replication	*NFIB*
ENSG00000147862	GO:0006351	transcription, DNA-templated	*NFIB*
ENSG00000147862	GO:0006355	regulation of transcription, DNA-templated	*NFIB*

ENSG00000147862	GO:0006366	transcription from RNA polymerase II promoter	*NFIB*
ENSG00000147862	GO:0008285	negative regulation of cell proliferation	*NFIB*
ENSG00000147862	GO:0010001	glial cell differentiation	*NFIB*
ENSG00000147862	GO:0021740	principal sensory nucleus of trigeminal nerve development	*NFIB*
ENSG00000147862	GO:0021960	anterior commissure morphogenesis	*NFIB*
ENSG00000147862	GO:0030324	lung development	*NFIB*
ENSG00000147862	GO:0030900	forebrain development	*NFIB*
ENSG00000147862	GO:0030902	hindbrain development	*NFIB*
ENSG00000147862	GO:0043392	negative regulation of DNA binding	*NFIB*
ENSG00000147862	GO:0044300	cerebellar mossy fiber	*NFIB*
ENSG00000147862	GO:0045944	positive regulation of transcription from RNA polymerase II promoter	*NFIB*
ENSG00000147862	GO:0060486	Clara cell differentiation	*NFIB*
ENSG00000147862	GO:0060509	Type I pneumocyte differentiation	*NFIB*
ENSG00000147862	GO:0060510	Type II pneumocyte differentiation	*NFIB*
ENSG00000147862	GO:0061141	lung ciliated cell differentiation	*NFIB*
ENSG00000147862	GO:0071679	commissural neuron axon guidance	*NFIB*
ENSG00000147862	GO:2000791	negative regulation of mesenchymal cell proliferation involved in lung development	*NFIB*
ENSG00000147862	GO:2000795	negative regulation of epithelial cell proliferation involved in lung morphogenesis	*NFIB*
ENSG00000139910	GO:0000398	mRNA splicing, via spliceosome	*NOVA1*
ENSG00000139910	GO:0003723	RNA binding	*NOVA1*
ENSG00000139910	GO:0003729	mRNA binding	*NOVA1*
ENSG00000139910	GO:0005634	nucleus	*NOVA1*
ENSG00000139910	GO:0005730	nucleolus	*NOVA1*
ENSG00000139910	GO:0006396	RNA processing	*NOVA1*
ENSG00000139910	GO:0007268	synaptic transmission	*NOVA1*
ENSG00000139910	GO:0007626	locomotory behavior	*NOVA1*
ENSG00000139910	GO:0008380	RNA splicing	*NOVA1*
ENSG00000139910	GO:0043231	intracellular membrane-bounded organelle	*NOVA1*
ENSG00000139910	GO:0044822	poly(A) RNA binding	*NOVA1*
ENSG00000139910	GO:0051252	regulation of RNA metabolic process	*NOVA1*
ENSG00000158458	GO:0005102	receptor binding	*NRG2*
ENSG00000158458	GO:0005515	protein binding	*NRG2*
ENSG00000158458	GO:0005576	extracellular region	*NRG2*
ENSG00000158458	GO:0005886	plasma membrane	*NRG2*
ENSG00000158458	GO:0007165	signal transduction	*NRG2*
ENSG00000158458	GO:0007173	epidermal growth factor receptor signaling pathway	*NRG2*
ENSG00000158458	GO:0008083	growth factor activity	*NRG2*
ENSG00000158458	GO:0008543	fibroblast growth factor receptor signaling pathway	*NRG2*
ENSG00000158458	GO:0009790	embryo development	*NRG2*
ENSG00000158458	GO:0016021	integral component of membrane	*NRG2*
ENSG00000158458	GO:0038095	Fc-epsilon receptor signaling pathway	*NRG2*
ENSG00000158458	GO:0045087	innate immune response	*NRG2*
ENSG00000158458	GO:0048011	neurotrophin TRK receptor signaling pathway	*NRG2*
ENSG00000158458	GO:0048015	phosphatidylinositol-mediated signaling	*NRG2*
ENSG00000144460	GO:0005515	protein binding	*NYAP2*

ENSG00000144460	GO:0014065	phosphatidylinositol 3-kinase signaling	*NYAP2*
ENSG00000144460	GO:0048812	neuron projection morphogenesis	*NYAP2*
ENSG00000168092	GO:0003847	1-alkyl-2-acetylglycerophosphocholine esterase activity	*PAFAH1B2*
ENSG00000168092	GO:0005515	protein binding	*PAFAH1B2*
ENSG00000168092	GO:0005730	nucleolus	*PAFAH1B2*
ENSG00000168092	GO:0005737	cytoplasm	*PAFAH1B2*
ENSG00000168092	GO:0005829	cytosol	*PAFAH1B2*
ENSG00000168092	GO:0005886	plasma membrane	*PAFAH1B2*
ENSG00000168092	GO:0006629	lipid metabolic process	*PAFAH1B2*
ENSG00000168092	GO:0007283	spermatogenesis	*PAFAH1B2*
ENSG00000168092	GO:0007420	brain development	*PAFAH1B2*
ENSG00000168092	GO:0008152	metabolic process	*PAFAH1B2*
ENSG00000168092	GO:0016042	lipid catabolic process	*PAFAH1B2*
ENSG00000168092	GO:0016239	positive regulation of macroautophagy	*PAFAH1B2*
ENSG00000168092	GO:0016787	hydrolase activity	*PAFAH1B2*
ENSG00000168092	GO:0016788	hydrolase activity, acting on ester bonds	*PAFAH1B2*
ENSG00000168092	GO:0042803	protein homodimerization activity	*PAFAH1B2*
ENSG00000168092	GO:0046982	protein heterodimerization activity	*PAFAH1B2*
ENSG00000168092	GO:0070062	extracellular vesicular exosome	*PAFAH1B2*
ENSG00000139946	GO:0001934	positive regulation of protein phosphorylation	*PELI2*
ENSG00000139946	GO:0002224	toll-like receptor signaling pathway	*PELI2*
ENSG00000139946	GO:0002755	MyD88-dependent toll-like receptor signaling pathway	*PELI2*
ENSG00000139946	GO:0005515	protein binding	*PELI2*
ENSG00000139946	GO:0005829	cytosol	*PELI2*
ENSG00000139946	GO:0008063	Toll signaling pathway	*PELI2*
ENSG00000139946	GO:0016567	protein ubiquitination	*PELI2*
ENSG00000139946	GO:0016874	ligase activity	*PELI2*
ENSG00000139946	GO:0034134	toll-like receptor 2 signaling pathway	*PELI2*
ENSG00000139946	GO:0034142	toll-like receptor 4 signaling pathway	*PELI2*
ENSG00000139946	GO:0034146	toll-like receptor 5 signaling pathway	*PELI2*
ENSG00000139946	GO:0034162	toll-like receptor 9 signaling pathway	*PELI2*
ENSG00000139946	GO:0034166	toll-like receptor 10 signaling pathway	*PELI2*
ENSG00000139946	GO:0038123	toll-like receptor TLR1:TLR2 signaling pathway	*PELI2*
ENSG00000139946	GO:0038124	toll-like receptor TLR6:TLR2 signaling pathway	*PELI2*
ENSG00000139946	GO:0043123	positive regulation of I-kappaB kinase/NF-kappaB signaling	*PELI2*
ENSG00000139946	GO:0043410	positive regulation of MAPK cascade	*PELI2*
ENSG00000139946	GO:0045087	innate immune response	*PELI2*
ENSG00000196570	GO:0003779	actin binding	*PFN3*
ENSG00000196570	GO:0005515	protein binding	*PFN3*
ENSG00000196570	GO:0005634	nucleus	*PFN3*
ENSG00000196570	GO:0005737	cytoplasm	*PFN3*
ENSG00000196570	GO:0005856	cytoskeleton	*PFN3*
ENSG00000196570	GO:0008289	lipid binding	*PFN3*
ENSG00000196570	GO:0030036	actin cytoskeleton organization	*PFN3*
ENSG00000112137	GO:0003779	actin binding	*PHACTR1*
ENSG00000112137	GO:0004864	protein phosphatase inhibitor activity	*PHACTR1*
ENSG00000112137	GO:0005634	nucleus	*PHACTR1*
ENSG00000112137	GO:0005829	cytosol	*PHACTR1*
ENSG00000112137	GO:0030054	cell junction	*PHACTR1*

ENSG00000112137	GO:0031032	actomyosin structure organization	*PHACTR1*
ENSG00000112137	GO:0031532	actin cytoskeleton reorganization	*PHACTR1*
ENSG00000112137	GO:0043086	negative regulation of catalytic activity	*PHACTR1*
ENSG00000112137	GO:0043149	stress fiber assembly	*PHACTR1*
ENSG00000112137	GO:0045202	synapse	*PHACTR1*
ENSG00000112137	GO:0048870	cell motility	*PHACTR1*
ENSG00000057294	GO:0002159	desmosome assembly	*PKP2*
ENSG00000057294	GO:0003674	molecular_function	*PKP2*
ENSG00000057294	GO:0005080	protein kinase C binding	*PKP2*
ENSG00000057294	GO:0005488	binding	*PKP2*
ENSG00000057294	GO:0005515	protein binding	*PKP2*
ENSG00000057294	GO:0005634	nucleus	*PKP2*
ENSG00000057294	GO:0005882	intermediate filament	*PKP2*
ENSG00000057294	GO:0005886	plasma membrane	*PKP2*
ENSG00000057294	GO:0005911	cell-cell junction	*PKP2*
ENSG00000057294	GO:0005912	adherens junction	*PKP2*
ENSG00000057294	GO:0007507	heart development	*PKP2*
ENSG00000057294	GO:0008285	negative regulation of cell proliferation	*PKP2*
ENSG00000057294	GO:0010765	positive regulation of sodium ion transport	*PKP2*
ENSG00000057294	GO:0014704	intercalated disc	*PKP2*
ENSG00000057294	GO:0016021	integral component of membrane	*PKP2*
ENSG00000057294	GO:0016264	gap junction assembly	*PKP2*
ENSG00000057294	GO:0016337	single organismal cell-cell adhesion	*PKP2*
ENSG00000057294	GO:0017080	sodium channel regulator activity	*PKP2*
ENSG00000057294	GO:0019215	intermediate filament binding	*PKP2*
ENSG00000057294	GO:0030054	cell junction	*PKP2*
ENSG00000057294	GO:0030057	desmosome	*PKP2*
ENSG00000057294	GO:0030336	negative regulation of cell migration	*PKP2*
ENSG00000057294	GO:0032947	protein complex scaffold	*PKP2*
ENSG00000057294	GO:0034334	adherens junction maintenance	*PKP2*
ENSG00000057294	GO:0044325	ion channel binding	*PKP2*
ENSG00000057294	GO:0045110	intermediate filament bundle assembly	*PKP2*
ENSG00000057294	GO:0048496	maintenance of organ identity	*PKP2*
ENSG00000057294	GO:0055010	ventricular cardiac muscle tissue morphogenesis	*PKP2*
ENSG00000057294	GO:0055088	lipid homeostasis	*PKP2*
ENSG00000057294	GO:0086001	cardiac muscle cell action potential	*PKP2*
ENSG00000057294	GO:0086002	cardiac muscle cell action potential involved in contraction	*PKP2*
ENSG00000057294	GO:0086005	ventricular cardiac muscle cell action potential	*PKP2*
ENSG00000057294	GO:0086019	cell-cell signaling involved in cardiac conduction	*PKP2*
ENSG00000057294	GO:0086064	cell communication by electrical coupling involved in cardiac conduction	*PKP2*
ENSG00000057294	GO:0086069	bundle of His cell to Purkinje myocyte communication	*PKP2*
ENSG00000057294	GO:0086091	regulation of heart rate by cardiac conduction	*PKP2*
ENSG00000057294	GO:2000810	regulation of tight junction assembly	*PKP2*
ENSG00000153246	GO:0001816	cytokine production	*PLA2R1*
ENSG00000153246	GO:0004872	receptor activity	*PLA2R1*
ENSG00000153246	GO:0005576	extracellular region	*PLA2R1*
ENSG00000153246	GO:0005886	plasma membrane	*PLA2R1*

ENSG00000153246	GO:0005887	integral component of plasma membrane	*PLA2R1*
ENSG00000153246	GO:0006898	receptor-mediated endocytosis	*PLA2R1*
ENSG00000153246	GO:0009986	cell surface	*PLA2R1*
ENSG00000153246	GO:0016021	integral component of membrane	*PLA2R1*
ENSG00000153246	GO:0030246	carbohydrate binding	*PLA2R1*
ENSG00000153246	GO:0043235	receptor complex	*PLA2R1*
ENSG00000153246	GO:0043274	phospholipase binding	*PLA2R1*
ENSG00000153246	GO:0043517	positive regulation of DNA damage response, signal transduction by p53 class mediator	*PLA2R1*
ENSG00000153246	GO:0070062	extracellular vesicular exosome	*PLA2R1*
ENSG00000153246	GO:0072593	reactive oxygen species metabolic process	*PLA2R1*
ENSG00000153246	GO:0090238	positive regulation of arachidonic acid secretion	*PLA2R1*
ENSG00000153246	GO:0090399	replicative senescence	*PLA2R1*
ENSG00000153246	GO:0090403	oxidative stress-induced premature senescence	*PLA2R1*
ENSG00000153246	GO:1900138	negative regulation of phospholipase A2 activity	*PLA2R1*
ENSG00000153246	GO:1900139	negative regulation of arachidonic acid secretion	*PLA2R1*
ENSG00000138814	GO:0000082	G1/S transition of mitotic cell cycle	*PPP3CA*
ENSG00000138814	GO:0001975	response to amphetamine	*PPP3CA*
ENSG00000138814	GO:0004721	phosphoprotein phosphatase activity	*PPP3CA*
ENSG00000138814	GO:0004722	protein serine/threonine phosphatase activity	*PPP3CA*
ENSG00000138814	GO:0004723	calcium-dependent protein serine/threonine phosphatase activity	*PPP3CA*
ENSG00000138814	GO:0005509	calcium ion binding	*PPP3CA*
ENSG00000138814	GO:0005515	protein binding	*PPP3CA*
ENSG00000138814	GO:0005516	calmodulin binding	*PPP3CA*
ENSG00000138814	GO:0005634	nucleus	*PPP3CA*
ENSG00000138814	GO:0005654	nucleoplasm	*PPP3CA*
ENSG00000138814	GO:0005737	cytoplasm	*PPP3CA*
ENSG00000138814	GO:0005739	mitochondrion	*PPP3CA*
ENSG00000138814	GO:0005829	cytosol	*PPP3CA*
ENSG00000138814	GO:0005955	calcineurin complex	*PPP3CA*
ENSG00000138814	GO:0006470	protein dephosphorylation	*PPP3CA*
ENSG00000138814	GO:0006606	protein import into nucleus	*PPP3CA*
ENSG00000138814	GO:0006816	calcium ion transport	*PPP3CA*
ENSG00000138814	GO:0008144	drug binding	*PPP3CA*
ENSG00000138814	GO:0014883	transition between fast and slow fiber	*PPP3CA*
ENSG00000138814	GO:0016020	membrane	*PPP3CA*
ENSG00000138814	GO:0016311	dephosphorylation	*PPP3CA*
ENSG00000138814	GO:0016787	hydrolase activity	*PPP3CA*
ENSG00000138814	GO:0019722	calcium-mediated signaling	*PPP3CA*
ENSG00000138814	GO:0019899	enzyme binding	*PPP3CA*
ENSG00000138814	GO:0030018	Z disc	*PPP3CA*
ENSG00000138814	GO:0033173	calcineurin-NFAT signaling cascade	*PPP3CA*
ENSG00000138814	GO:0033192	calmodulin-dependent protein phosphatase activity	*PPP3CA*
ENSG00000138814	GO:0033555	multicellular organismal response to stress	*PPP3CA*
ENSG00000138814	GO:0035690	cellular response to drug	*PPP3CA*
ENSG00000138814	GO:0038095	Fc-epsilon receptor signaling pathway	*PPP3CA*

ENSG00000138814	GO:0042110	T cell activation	*PPP3CA*
ENSG00000138814	GO:0042383	sarcolemma	*PPP3CA*
ENSG00000138814	GO:0045087	innate immune response	*PPP3CA*
ENSG00000138814	GO:0045944	positive regulation of transcription from RNA polymerase II promoter	*PPP3CA*
ENSG00000138814	GO:0046676	negative regulation of insulin secretion	*PPP3CA*
ENSG00000138814	GO:0046982	protein heterodimerization activity	*PPP3CA*
ENSG00000138814	GO:0046983	protein dimerization activity	*PPP3CA*
ENSG00000138814	GO:0048741	skeletal muscle fiber development	*PPP3CA*
ENSG00000138814	GO:0050804	regulation of synaptic transmission	*PPP3CA*
ENSG00000138814	GO:0051592	response to calcium ion	*PPP3CA*
ENSG00000138814	GO:0060079	regulation of excitatory postsynaptic membrane potential	*PPP3CA*
ENSG00000138814	GO:0071333	cellular response to glucose stimulus	*PPP3CA*
ENSG00000122420	GO:0004930	G-protein coupled receptor activity	*PTGFR*
ENSG00000122420	GO:0004958	prostaglandin F receptor activity	*PTGFR*
ENSG00000122420	GO:0005576	extracellular region	*PTGFR*
ENSG00000122420	GO:0005737	cytoplasm	*PTGFR*
ENSG00000122420	GO:0005886	plasma membrane	*PTGFR*
ENSG00000122420	GO:0005887	integral component of plasma membrane	*PTGFR*
ENSG00000122420	GO:0007186	G-protein coupled receptor signaling pathway	*PTGFR*
ENSG00000122420	GO:0007567	parturition	*PTGFR*
ENSG00000122420	GO:0016021	integral component of membrane	*PTGFR*
ENSG00000122420	GO:0032496	response to lipopolysaccharide	*PTGFR*
ENSG00000122420	GO:0035584	calcium-mediated signaling using intracellular calcium source	*PTGFR*
ENSG00000122420	GO:0043066	negative regulation of apoptotic process	*PTGFR*
ENSG00000144642	GO:0000166	nucleotide binding	*RBMS3*
ENSG00000144642	GO:0003676	nucleic acid binding	*RBMS3*
ENSG00000144642	GO:0003723	RNA binding	*RBMS3*
ENSG00000144642	GO:0003730	mRNA 3'-UTR binding	*RBMS3*
ENSG00000144642	GO:0005737	cytoplasm	*RBMS3*
ENSG00000144642	GO:0045727	positive regulation of translation	*RBMS3*
ENSG00000148908	GO:0005096	GTPase activator activity	*RGS10*
ENSG00000148908	GO:0005515	protein binding	*RGS10*
ENSG00000148908	GO:0005634	nucleus	*RGS10*
ENSG00000148908	GO:0005737	cytoplasm	*RGS10*
ENSG00000148908	GO:0005886	plasma membrane	*RGS10*
ENSG00000148908	GO:0008277	regulation of G-protein coupled receptor protein signaling pathway	*RGS10*
ENSG00000148908	GO:0038032	termination of G-protein coupled receptor signaling pathway	*RGS10*
ENSG00000148908	GO:0043025	neuronal cell body	*RGS10*
ENSG00000148908	GO:0043197	dendritic spine	*RGS10*
ENSG00000148908	GO:0043547	positive regulation of GTPase activity	*RGS10*
ENSG00000148908	GO:0043679	axon terminus	*RGS10*
ENSG00000169220	GO:0000922	spindle pole	*RGS14*
ENSG00000169220	GO:0001965	G-protein alpha-subunit binding	*RGS14*
ENSG00000169220	GO:0005057	receptor signaling protein activity	*RGS14*
ENSG00000169220	GO:0005092	GDP-dissociation inhibitor activity	*RGS14*
ENSG00000169220	GO:0005096	GTPase activator activity	*RGS14*
ENSG00000169220	GO:0005515	protein binding	*RGS14*
ENSG00000169220	GO:0005634	nucleus	*RGS14*

ENSG00000169220	GO:0005737	cytoplasm	*RGS14*
ENSG00000169220	GO:0005813	centrosome	*RGS14*
ENSG00000169220	GO:0005819	spindle	*RGS14*
ENSG00000169220	GO:0005874	microtubule	*RGS14*
ENSG00000169220	GO:0005886	plasma membrane	*RGS14*
ENSG00000169220	GO:0006913	nucleocytoplasmic transport	*RGS14*
ENSG00000169220	GO:0006979	response to oxidative stress	*RGS14*
ENSG00000169220	GO:0007051	spindle organization	*RGS14*
ENSG00000169220	GO:0007059	chromosome segregation	*RGS14*
ENSG00000169220	GO:0007067	mitotic nuclear division	*RGS14*
ENSG00000169220	GO:0007165	signal transduction	*RGS14*
ENSG00000169220	GO:0007612	learning	*RGS14*
ENSG00000169220	GO:0007616	long-term memory	*RGS14*
ENSG00000169220	GO:0008017	microtubule binding	*RGS14*
ENSG00000169220	GO:0008277	regulation of G-protein coupled receptor protein signaling pathway	*RGS14*
ENSG00000169220	GO:0008542	visual learning	*RGS14*
ENSG00000169220	GO:0010070	zygote asymmetric cell division	*RGS14*
ENSG00000169220	GO:0014069	postsynaptic density	*RGS14*
ENSG00000169220	GO:0016604	nuclear body	*RGS14*
ENSG00000169220	GO:0016605	PML body	*RGS14*
ENSG00000169220	GO:0019901	protein kinase binding	*RGS14*
ENSG00000169220	GO:0030054	cell junction	*RGS14*
ENSG00000169220	GO:0030159	receptor signaling complex scaffold activity	*RGS14*
ENSG00000169220	GO:0030425	dendrite	*RGS14*
ENSG00000169220	GO:0030695	GTPase regulator activity	*RGS14*
ENSG00000169220	GO:0031914	negative regulation of synaptic plasticity	*RGS14*
ENSG00000169220	GO:0032794	GTPase activating protein binding	*RGS14*
ENSG00000169220	GO:0035556	intracellular signal transduction	*RGS14*
ENSG00000169220	GO:0038032	termination of G-protein coupled receptor signaling pathway	*RGS14*
ENSG00000169220	GO:0043197	dendritic spine	*RGS14*
ENSG00000169220	GO:0043407	negative regulation of MAP kinase activity	*RGS14*
ENSG00000169220	GO:0043547	positive regulation of GTPase activity	*RGS14*
ENSG00000169220	GO:0043620	regulation of DNA-templated transcription in response to stress	*RGS14*
ENSG00000169220	GO:0045211	postsynaptic membrane	*RGS14*
ENSG00000169220	GO:0048008	platelet-derived growth factor receptor signaling pathway	*RGS14*
ENSG00000169220	GO:0050769	positive regulation of neurogenesis	*RGS14*
ENSG00000169220	GO:0051301	cell division	*RGS14*
ENSG00000169220	GO:0060291	long-term synaptic potentiation	*RGS14*
ENSG00000169220	GO:0070373	negative regulation of ERK1 and ERK2 cascade	*RGS14*
ENSG00000069667	GO:0001222	transcription corepressor binding	*RORA*
ENSG00000069667	GO:0001223	transcription coactivator binding	*RORA*
ENSG00000069667	GO:0001525	angiogenesis	*RORA*
ENSG00000069667	GO:0003677	DNA binding	*RORA*
ENSG00000069667	GO:0003700	sequence-specific DNA binding transcription factor activity	*RORA*
ENSG00000069667	GO:0003707	steroid hormone receptor activity	*RORA*
ENSG00000069667	GO:0004879	ligand-activated sequence-specific DNA binding RNA polymerase II transcription factor activity	*RORA*

ENSG00000069667	GO:0005515	protein binding	*RORA*
ENSG00000069667	GO:0005634	nucleus	*RORA*
ENSG00000069667	GO:0005654	nucleoplasm	*RORA*
ENSG00000069667	GO:0006351	transcription, DNA-templated	*RORA*
ENSG00000069667	GO:0006355	regulation of transcription, DNA-templated	*RORA*
ENSG00000069667	GO:0006367	transcription initiation from RNA polymerase II promoter	*RORA*
ENSG00000069667	GO:0006805	xenobiotic metabolic process	*RORA*
ENSG00000069667	GO:0006809	nitric oxide biosynthetic process	*RORA*
ENSG00000069667	GO:0007623	circadian rhythm	*RORA*
ENSG00000069667	GO:0008013	beta-catenin binding	*RORA*
ENSG00000069667	GO:0008134	transcription factor binding	*RORA*
ENSG00000069667	GO:0008142	oxysterol binding	*RORA*
ENSG00000069667	GO:0008270	zinc ion binding	*RORA*
ENSG00000069667	GO:0008589	regulation of smoothened signaling pathway	*RORA*
ENSG00000069667	GO:0010467	gene expression	*RORA*
ENSG00000069667	GO:0010575	positive regulation vascular endothelial growth factor production	*RORA*
ENSG00000069667	GO:0010906	regulation of glucose metabolic process	*RORA*
ENSG00000069667	GO:0019218	regulation of steroid metabolic process	*RORA*
ENSG00000069667	GO:0021702	cerebellar Purkinje cell differentiation	*RORA*
ENSG00000069667	GO:0021930	cerebellar granule cell precursor proliferation	*RORA*
ENSG00000069667	GO:0030522	intracellular receptor signaling pathway	*RORA*
ENSG00000069667	GO:0032922	circadian regulation of gene expression	*RORA*
ENSG00000069667	GO:0036315	cellular response to sterol	*RORA*
ENSG00000069667	GO:0042692	muscle cell differentiation	*RORA*
ENSG00000069667	GO:0042752	regulation of circadian rhythm	*RORA*
ENSG00000069667	GO:0042753	positive regulation of circadian rhythm	*RORA*
ENSG00000069667	GO:0043030	regulation of macrophage activation	*RORA*
ENSG00000069667	GO:0043124	negative regulation of I-kappaB kinase/NF-kappaB signaling	*RORA*
ENSG00000069667	GO:0043401	steroid hormone mediated signaling pathway	*RORA*
ENSG00000069667	GO:0043565	sequence-specific DNA binding	*RORA*
ENSG00000069667	GO:0045599	negative regulation of fat cell differentiation	*RORA*
ENSG00000069667	GO:0045893	positive regulation of transcription, DNA-templated	*RORA*
ENSG00000069667	GO:0045944	positive regulation of transcription from RNA polymerase II promoter	*RORA*
ENSG00000069667	GO:0046068	cGMP metabolic process	*RORA*
ENSG00000069667	GO:0050728	negative regulation of inflammatory response	*RORA*
ENSG00000069667	GO:0060850	regulation of transcription involved in cell fate commitment	*RORA*
ENSG00000069667	GO:0070328	triglyceride homeostasis	*RORA*
ENSG00000069667	GO:0071456	cellular response to hypoxia	*RORA*
ENSG00000069667	GO:0072539	T-helper 17 cell differentiation	*RORA*
ENSG00000069667	GO:0098531	direct ligand regulated sequence-specific DNA binding transcription factor activity	*RORA*

ENSG00000069667	GO:2000188	regulation of cholesterol homeostasis	*RORA*
ENSG00000137872	GO:0004872	receptor activity	*SEMA6D*
ENSG00000137872	GO:0005515	protein binding	*SEMA6D*
ENSG00000137872	GO:0005737	cytoplasm	*SEMA6D*
ENSG00000137872	GO:0005886	plasma membrane	*SEMA6D*
ENSG00000137872	GO:0007275	multicellular organismal development	*SEMA6D*
ENSG00000137872	GO:0007411	axon guidance	*SEMA6D*
ENSG00000137872	GO:0014911	positive regulation of smooth muscle cell migration	*SEMA6D*
ENSG00000137872	GO:0014912	negative regulation of smooth muscle cell migration	*SEMA6D*
ENSG00000137872	GO:0016020	membrane	*SEMA6D*
ENSG00000137872	GO:0016021	integral component of membrane	*SEMA6D*
ENSG00000137872	GO:0021591	ventricular system development	*SEMA6D*
ENSG00000137872	GO:0030215	semaphorin receptor binding	*SEMA6D*
ENSG00000198964	GO:0000138	Golgi trans cisterna	*SGMS1*
ENSG00000198964	GO:0000139	Golgi membrane	*SGMS1*
ENSG00000198964	GO:0003824	catalytic activity	*SGMS1*
ENSG00000198964	GO:0005515	protein binding	*SGMS1*
ENSG00000198964	GO:0005634	nucleus	*SGMS1*
ENSG00000198964	GO:0005783	endoplasmic reticulum	*SGMS1*
ENSG00000198964	GO:0005886	plasma membrane	*SGMS1*
ENSG00000198964	GO:0006665	sphingolipid metabolic process	*SGMS1*
ENSG00000198964	GO:0006686	sphingomyelin biosynthetic process	*SGMS1*
ENSG00000198964	GO:0006915	apoptotic process	*SGMS1*
ENSG00000198964	GO:0006954	inflammatory response	*SGMS1*
ENSG00000198964	GO:0010628	positive regulation of gene expression	*SGMS1*
ENSG00000198964	GO:0016020	membrane	*SGMS1*
ENSG00000198964	GO:0016021	integral component of membrane	*SGMS1*
ENSG00000198964	GO:0016049	cell growth	*SGMS1*
ENSG00000198964	GO:0016301	kinase activity	*SGMS1*
ENSG00000198964	GO:0016310	phosphorylation	*SGMS1*
ENSG00000198964	GO:0030148	sphingolipid biosynthetic process	*SGMS1*
ENSG00000198964	GO:0030173	integral component of Golgi membrane	*SGMS1*
ENSG00000198964	GO:0033188	sphingomyelin synthase activity	*SGMS1*
ENSG00000198964	GO:0044281	small molecule metabolic process	*SGMS1*
ENSG00000198964	GO:0047493	ceramide cholinephosphotransferase activity	*SGMS1*
ENSG00000198964	GO:0071222	cellular response to lipopolysaccharide	*SGMS1*
ENSG00000198964	GO:0071356	cellular response to tumor necrosis factor	*SGMS1*
ENSG00000198964	GO:2001237	negative regulation of extrinsic apoptotic signaling pathway	*SGMS1*
ENSG00000116991	GO:0005096	GTPase activator activity	*SIPA1L2*
ENSG00000116991	GO:0005515	protein binding	*SIPA1L2*
ENSG00000116991	GO:0043547	positive regulation of GTPase activity	*SIPA1L2*
ENSG00000116991	GO:0051056	regulation of small GTPase mediated signal transduction	*SIPA1L2*
ENSG00000180251	GO:0001696	gastric acid secretion	*SLC9A4*
ENSG00000180251	GO:0002064	epithelial cell development	*SLC9A4*
ENSG00000180251	GO:0005886	plasma membrane	*SLC9A4*
ENSG00000180251	GO:0006811	ion transport	*SLC9A4*
ENSG00000180251	GO:0006812	cation transport	*SLC9A4*
ENSG00000180251	GO:0006814	sodium ion transport	*SLC9A4*

ENSG00000180251	GO:0006885	regulation of pH	*SLC9A4*
ENSG00000180251	GO:0015299	solute:proton antiporter activity	*SLC9A4*
ENSG00000180251	GO:0015385	sodium:proton antiporter activity	*SLC9A4*
ENSG00000180251	GO:0016021	integral component of membrane	*SLC9A4*
ENSG00000180251	GO:0016323	basolateral plasma membrane	*SLC9A4*
ENSG00000180251	GO:0016324	apical plasma membrane	*SLC9A4*
ENSG00000180251	GO:0035725	sodium ion transmembrane transport	*SLC9A4*
ENSG00000180251	GO:0055085	transmembrane transport	*SLC9A4*
ENSG00000180251	GO:1902600	hydrogen ion transmembrane transport	*SLC9A4*
ENSG00000125398	GO:0000976	transcription regulatory region sequence-specific DNA binding	*SOX9*
ENSG00000125398	GO:0000981	sequence-specific DNA binding RNA polymerase II transcription factor activity	*SOX9*
ENSG00000125398	GO:0001046	core promoter sequence-specific DNA binding	*SOX9*
ENSG00000125398	GO:0001077	RNA polymerase II core promoter proximal region sequence-specific DNA binding transcription factor activity involved in positive regulation of transcription	*SOX9*
ENSG00000125398	GO:0001158	enhancer sequence-specific DNA binding	*SOX9*
ENSG00000125398	GO:0001501	skeletal system development	*SOX9*
ENSG00000125398	GO:0001502	cartilage condensation	*SOX9*
ENSG00000125398	GO:0001503	ossification	*SOX9*
ENSG00000125398	GO:0001658	branching involved in ureteric bud morphogenesis	*SOX9*
ENSG00000125398	GO:0001708	cell fate specification	*SOX9*
ENSG00000125398	GO:0001837	epithelial to mesenchymal transition	*SOX9*
ENSG00000125398	GO:0001894	tissue homeostasis	*SOX9*
ENSG00000125398	GO:0001934	positive regulation of protein phosphorylation	*SOX9*
ENSG00000125398	GO:0001942	hair follicle development	*SOX9*
ENSG00000125398	GO:0002053	positive regulation of mesenchymal cell proliferation	*SOX9*
ENSG00000125398	GO:0002062	chondrocyte differentiation	*SOX9*
ENSG00000125398	GO:0002063	chondrocyte development	*SOX9*
ENSG00000125398	GO:0002683	negative regulation of immune system process	*SOX9*
ENSG00000125398	GO:0003170	heart valve development	*SOX9*
ENSG00000125398	GO:0003179	heart valve morphogenesis	*SOX9*
ENSG00000125398	GO:0003188	heart valve formation	*SOX9*
ENSG00000125398	GO:0003203	endocardial cushion morphogenesis	*SOX9*
ENSG00000125398	GO:0003413	chondrocyte differentiation involved in endochondral bone morphogenesis	*SOX9*
ENSG00000125398	GO:0003415	chondrocyte hypertrophy	*SOX9*
ENSG00000125398	GO:0003677	DNA binding	*SOX9*
ENSG00000125398	GO:0003682	chromatin binding	*SOX9*
ENSG00000125398	GO:0003700	sequence-specific DNA binding transcription factor activity	*SOX9*
ENSG00000125398	GO:0003705	RNA polymerase II distal enhancer sequence-specific DNA binding transcription factor activity	*SOX9*
ENSG00000125398	GO:0004672	protein kinase activity	*SOX9*
ENSG00000125398	GO:0005515	protein binding	*SOX9*
ENSG00000125398	GO:0005634	nucleus	*SOX9*

ENSG00000125398	GO:0005667	transcription factor complex	*SOX9*
ENSG00000125398	GO:0006334	nucleosome assembly	*SOX9*
ENSG00000125398	GO:0006338	chromatin remodeling	*SOX9*
ENSG00000125398	GO:0006357	regulation of transcription from RNA polymerase II promoter	*SOX9*
ENSG00000125398	GO:0006366	transcription from RNA polymerase II promoter	*SOX9*
ENSG00000125398	GO:0006461	protein complex assembly	*SOX9*
ENSG00000125398	GO:0006468	protein phosphorylation	*SOX9*
ENSG00000125398	GO:0007010	cytoskeleton organization	*SOX9*
ENSG00000125398	GO:0007165	signal transduction	*SOX9*
ENSG00000125398	GO:0007173	epidermal growth factor receptor signaling pathway	*SOX9*
ENSG00000125398	GO:0007283	spermatogenesis	*SOX9*
ENSG00000125398	GO:0007417	central nervous system development	*SOX9*
ENSG00000125398	GO:0007507	heart development	*SOX9*
ENSG00000125398	GO:0008013	beta-catenin binding	*SOX9*
ENSG00000125398	GO:0008284	positive regulation of cell proliferation	*SOX9*
ENSG00000125398	GO:0008285	negative regulation of cell proliferation	*SOX9*
ENSG00000125398	GO:0008584	male gonad development	*SOX9*
ENSG00000125398	GO:0010564	regulation of cell cycle process	*SOX9*
ENSG00000125398	GO:0010628	positive regulation of gene expression	*SOX9*
ENSG00000125398	GO:0010634	positive regulation of epithelial cell migration	*SOX9*
ENSG00000125398	GO:0014032	neural crest cell development	*SOX9*
ENSG00000125398	GO:0014068	positive regulation of phosphatidylinositol 3-kinase signaling	*SOX9*
ENSG00000125398	GO:0016337	single organismal cell-cell adhesion	*SOX9*
ENSG00000125398	GO:0019100	male germ-line sex determination	*SOX9*
ENSG00000125398	GO:0019933	cAMP-mediated signaling	*SOX9*
ENSG00000125398	GO:0030154	cell differentiation	*SOX9*
ENSG00000125398	GO:0030155	regulation of cell adhesion	*SOX9*
ENSG00000125398	GO:0030198	extracellular matrix organization	*SOX9*
ENSG00000125398	GO:0030238	male sex determination	*SOX9*
ENSG00000125398	GO:0030279	negative regulation of ossification	*SOX9*
ENSG00000125398	GO:0030502	negative regulation of bone mineralization	*SOX9*
ENSG00000125398	GO:0030850	prostate gland development	*SOX9*
ENSG00000125398	GO:0030857	negative regulation of epithelial cell differentiation	*SOX9*
ENSG00000125398	GO:0030858	positive regulation of epithelial cell differentiation	*SOX9*
ENSG00000125398	GO:0030879	mammary gland development	*SOX9*
ENSG00000125398	GO:0030903	notochord development	*SOX9*
ENSG00000125398	GO:0030916	otic vesicle formation	*SOX9*
ENSG00000125398	GO:0031018	endocrine pancreas development	*SOX9*
ENSG00000125398	GO:0032331	negative regulation of chondrocyte differentiation	*SOX9*
ENSG00000125398	GO:0032332	positive regulation of chondrocyte differentiation	*SOX9*
ENSG00000125398	GO:0034236	protein kinase A catalytic subunit binding	*SOX9*
ENSG00000125398	GO:0035019	somatic stem cell maintenance	*SOX9*
ENSG00000125398	GO:0035326	enhancer binding	*SOX9*
ENSG00000125398	GO:0035622	intrahepatic bile duct development	*SOX9*
ENSG00000125398	GO:0042127	regulation of cell proliferation	*SOX9*
ENSG00000125398	GO:0042981	regulation of apoptotic process	*SOX9*

ENSG00000125398	GO:0043066	negative regulation of apoptotic process	*SOX9*
ENSG00000125398	GO:0043234	protein complex	*SOX9*
ENSG00000125398	GO:0043425	bHLH transcription factor binding	*SOX9*
ENSG00000125398	GO:0043491	protein kinase B signaling	*SOX9*
ENSG00000125398	GO:0043565	sequence-specific DNA binding	*SOX9*
ENSG00000125398	GO:0044212	transcription regulatory region DNA binding	*SOX9*
ENSG00000125398	GO:0044798	nuclear transcription factor complex	*SOX9*
ENSG00000125398	GO:0045165	cell fate commitment	*SOX9*
ENSG00000125398	GO:0045595	regulation of cell differentiation	*SOX9*
ENSG00000125398	GO:0045662	negative regulation of myoblast differentiation	*SOX9*
ENSG00000125398	GO:0045732	positive regulation of protein catabolic process	*SOX9*
ENSG00000125398	GO:0045892	negative regulation of transcription, DNA-templated	*SOX9*
ENSG00000125398	GO:0045893	positive regulation of transcription, DNA-templated	*SOX9*
ENSG00000125398	GO:0045944	positive regulation of transcription from RNA polymerase II promoter	*SOX9*
ENSG00000125398	GO:0046533	negative regulation of photoreceptor cell differentiation	*SOX9*
ENSG00000125398	GO:0046982	protein heterodimerization activity	*SOX9*
ENSG00000125398	GO:0048709	oligodendrocyte differentiation	*SOX9*
ENSG00000125398	GO:0050679	positive regulation of epithelial cell proliferation	*SOX9*
ENSG00000125398	GO:0050680	negative regulation of epithelial cell proliferation	*SOX9*
ENSG00000125398	GO:0051216	cartilage development	*SOX9*
ENSG00000125398	GO:0060008	Sertoli cell differentiation	*SOX9*
ENSG00000125398	GO:0060009	Sertoli cell development	*SOX9*
ENSG00000125398	GO:0060018	astrocyte fate commitment	*SOX9*
ENSG00000125398	GO:0060041	retina development in camera-type eye	*SOX9*
ENSG00000125398	GO:0060174	limb bud formation	*SOX9*
ENSG00000125398	GO:0060221	retinal rod cell differentiation	*SOX9*
ENSG00000125398	GO:0060350	endochondral bone morphogenesis	*SOX9*
ENSG00000125398	GO:0060441	epithelial tube branching involved in lung morphogenesis	*SOX9*
ENSG00000125398	GO:0060487	lung epithelial cell differentiation	*SOX9*
ENSG00000125398	GO:0060512	prostate gland morphogenesis	*SOX9*
ENSG00000125398	GO:0060517	epithelial cell proliferation involved in prostatic bud elongation	*SOX9*
ENSG00000125398	GO:0060729	intestinal epithelial structure maintenance	*SOX9*
ENSG00000125398	GO:0060784	regulation of cell proliferation involved in tissue homeostasis	*SOX9*
ENSG00000125398	GO:0061036	positive regulation of cartilage development	*SOX9*
ENSG00000125398	GO:0061046	regulation of branching involved in lung morphogenesis	*SOX9*
ENSG00000125398	GO:0061138	morphogenesis of a branching epithelium	*SOX9*
ENSG00000125398	GO:0070168	negative regulation of biomineral tissue development	*SOX9*
ENSG00000125398	GO:0070371	ERK1 and ERK2 cascade	*SOX9*
ENSG00000125398	GO:0071260	cellular response to mechanical	*SOX9*

		stimulus	
ENSG00000125398	GO:0071300	cellular response to retinoic acid	*SOX9*
ENSG00000125398	GO:0071347	cellular response to interleukin-1	*SOX9*
ENSG00000125398	GO:0071364	cellular response to epidermal growth factor stimulus	*SOX9*
ENSG00000125398	GO:0071504	cellular response to heparin	*SOX9*
ENSG00000125398	GO:0071560	cellular response to transforming growth factor beta stimulus	*SOX9*
ENSG00000125398	GO:0071599	otic vesicle development	*SOX9*
ENSG00000125398	GO:0072034	renal vesicle induction	*SOX9*
ENSG00000125398	GO:0072170	metanephric tubule development	*SOX9*
ENSG00000125398	GO:0072189	ureter development	*SOX9*
ENSG00000125398	GO:0072190	ureter urothelium development	*SOX9*
ENSG00000125398	GO:0072193	ureter smooth muscle cell differentiation	*SOX9*
ENSG00000125398	GO:0072197	ureter morphogenesis	*SOX9*
ENSG00000125398	GO:0072289	metanephric nephron tubule formation	*SOX9*
ENSG00000125398	GO:0090090	negative regulation of canonical Wnt signaling pathway	*SOX9*
ENSG00000125398	GO:0090103	cochlea morphogenesis	*SOX9*
ENSG00000125398	GO:0090184	positive regulation of kidney development	*SOX9*
ENSG00000125398	GO:0090190	positive regulation of branching involved in ureteric bud morphogenesis	*SOX9*
ENSG00000125398	GO:0097157	pre-mRNA intronic binding	*SOX9*
ENSG00000125398	GO:1901203	positive regulation of extracellular matrix assembly	*SOX9*
ENSG00000125398	GO:2000020	positive regulation of male gonad development	*SOX9*
ENSG00000125398	GO:2000138	positive regulation of cell proliferation involved in heart morphogenesis	*SOX9*
ENSG00000125398	GO:2000741	positive regulation of mesenchymal stem cell differentiation	*SOX9*
ENSG00000125398	GO:2000794	regulation of epithelial cell proliferation involved in lung morphogenesis	*SOX9*
ENSG00000125398	GO:2001054	negative regulation of mesenchymal cell apoptotic process	*SOX9*
ENSG00000164056	GO:0000132	establishment of mitotic spindle orientation	*SPRY1*
ENSG00000164056	GO:0001656	metanephros development	*SPRY1*
ENSG00000164056	GO:0001657	ureteric bud development	*SPRY1*
ENSG00000164056	GO:0001759	organ induction	*SPRY1*
ENSG00000164056	GO:0005515	protein binding	*SPRY1*
ENSG00000164056	GO:0005829	cytosol	*SPRY1*
ENSG00000164056	GO:0005886	plasma membrane	*SPRY1*
ENSG00000164056	GO:0007173	epidermal growth factor receptor signaling pathway	*SPRY1*
ENSG00000164056	GO:0007275	multicellular organismal development	*SPRY1*
ENSG00000164056	GO:0008285	negative regulation of cell proliferation	*SPRY1*
ENSG00000164056	GO:0009966	regulation of signal transduction	*SPRY1*
ENSG00000164056	GO:0016020	membrane	*SPRY1*
ENSG00000164056	GO:0034261	negative regulation of Ras GTPase activity	*SPRY1*
ENSG00000164056	GO:0040037	negative regulation of fibroblast growth factor receptor signaling pathway	*SPRY1*
ENSG00000164056	GO:0042059	negative regulation of epidermal	*SPRY1*

		growth factor receptor signaling pathway	
ENSG00000164056	GO:0043407	negative regulation of MAP kinase activity	*SPRY1*
ENSG00000164056	GO:0046580	negative regulation of Ras protein signal transduction	*SPRY1*
ENSG00000164056	GO:0051387	negative regulation of neurotrophin TRK receptor signaling pathway	*SPRY1*
ENSG00000164056	GO:0060449	bud elongation involved in lung branching	*SPRY1*
ENSG00000164056	GO:0070373	negative regulation of ERK1 and ERK2 cascade	*SPRY1*
ENSG00000114850	GO:0005784	Sec61 translocon complex	*SSR3*
ENSG00000114850	GO:0006412	translation	*SSR3*
ENSG00000114850	GO:0006613	cotranslational protein targeting to membrane	*SSR3*
ENSG00000114850	GO:0006614	SRP-dependent cotranslational protein targeting to membrane	*SSR3*
ENSG00000114850	GO:0010467	gene expression	*SSR3*
ENSG00000114850	GO:0030176	integral component of endoplasmic reticulum membrane	*SSR3*
ENSG00000114850	GO:0044267	cellular protein metabolic process	*SSR3*
ENSG00000068781	GO:0005515	protein binding	*STON1-GTF2A1L*
ENSG00000068781	GO:0005672	transcription factor TFIIA complex	*STON1-GTF2A1L*
ENSG00000068781	GO:0006367	transcription initiation from RNA polymerase II promoter	*STON1-GTF2A1L*
ENSG00000068781	GO:0006886	intracellular protein transport	*STON1-GTF2A1L*
ENSG00000068781	GO:0006897	endocytosis	*STON1-GTF2A1L*
ENSG00000068781	GO:0016192	vesicle-mediated transport	*STON1-GTF2A1L*
ENSG00000068781	GO:0030131	clathrin adaptor complex	*STON1-GTF2A1L*
ENSG00000163513	GO:0001568	blood vessel development	*TGFBR2*
ENSG00000163513	GO:0001569	patterning of blood vessels	*TGFBR2*
ENSG00000163513	GO:0001570	vasculogenesis	*TGFBR2*
ENSG00000163513	GO:0001701	in utero embryonic development	*TGFBR2*
ENSG00000163513	GO:0002053	positive regulation of mesenchymal cell proliferation	*TGFBR2*
ENSG00000163513	GO:0002088	lens development in camera-type eye	*TGFBR2*
ENSG00000163513	GO:0002651	positive regulation of tolerance induction to self antigen	*TGFBR2*
ENSG00000163513	GO:0002663	positive regulation of B cell tolerance induction	*TGFBR2*
ENSG00000163513	GO:0002666	positive regulation of T cell tolerance induction	*TGFBR2*
ENSG00000163513	GO:0004672	protein kinase activity	*TGFBR2*
ENSG00000163513	GO:0004674	protein serine/threonine kinase activity	*TGFBR2*
ENSG00000163513	GO:0004675	transmembrane receptor protein serine/threonine kinase activity	*TGFBR2*
ENSG00000163513	GO:0004702	receptor signaling protein serine/threonine kinase activity	*TGFBR2*
ENSG00000163513	GO:0004713	protein tyrosine kinase activity	*TGFBR2*
ENSG00000163513	GO:0004872	receptor activity	*TGFBR2*
ENSG00000163513	GO:0005024	transforming growth factor beta-	*TGFBR2*

		activated receptor activity	
ENSG00000163513	GO:0005026	transforming growth factor beta receptor activity, type II	*TGFBR2*
ENSG00000163513	GO:0005515	protein binding	*TGFBR2*
ENSG00000163513	GO:0005524	ATP binding	*TGFBR2*
ENSG00000163513	GO:0005539	glycosaminoglycan binding	*TGFBR2*
ENSG00000163513	GO:0005737	cytoplasm	*TGFBR2*
ENSG00000163513	GO:0005829	cytosol	*TGFBR2*
ENSG00000163513	GO:0005886	plasma membrane	*TGFBR2*
ENSG00000163513	GO:0005887	integral component of plasma membrane	*TGFBR2*
ENSG00000163513	GO:0005901	caveola	*TGFBR2*
ENSG00000163513	GO:0006468	protein phosphorylation	*TGFBR2*
ENSG00000163513	GO:0006898	receptor-mediated endocytosis	*TGFBR2*
ENSG00000163513	GO:0006915	apoptotic process	*TGFBR2*
ENSG00000163513	GO:0007179	transforming growth factor beta receptor signaling pathway	*TGFBR2*
ENSG00000163513	GO:0007182	common-partner SMAD protein phosphorylation	*TGFBR2*
ENSG00000163513	GO:0007224	smoothened signaling pathway	*TGFBR2*
ENSG00000163513	GO:0007369	gastrulation	*TGFBR2*
ENSG00000163513	GO:0007420	brain development	*TGFBR2*
ENSG00000163513	GO:0007507	heart development	*TGFBR2*
ENSG00000163513	GO:0007566	embryo implantation	*TGFBR2*
ENSG00000163513	GO:0007568	aging	*TGFBR2*
ENSG00000163513	GO:0007584	response to nutrient	*TGFBR2*
ENSG00000163513	GO:0008284	positive regulation of cell proliferation	*TGFBR2*
ENSG00000163513	GO:0008285	negative regulation of cell proliferation	*TGFBR2*
ENSG00000163513	GO:0009612	response to mechanical stimulus	*TGFBR2*
ENSG00000163513	GO:0009749	response to glucose	*TGFBR2*
ENSG00000163513	GO:0009887	organ morphogenesis	*TGFBR2*
ENSG00000163513	GO:0009897	external side of plasma membrane	*TGFBR2*
ENSG00000163513	GO:0009986	cell surface	*TGFBR2*
ENSG00000163513	GO:0010033	response to organic substance	*TGFBR2*
ENSG00000163513	GO:0010634	positive regulation of epithelial cell migration	*TGFBR2*
ENSG00000163513	GO:0014070	response to organic cyclic compound	*TGFBR2*
ENSG00000163513	GO:0016020	membrane	*TGFBR2*
ENSG00000163513	GO:0016021	integral component of membrane	*TGFBR2*
ENSG00000163513	GO:0016772	transferase activity, transferring phosphorus-containing groups	*TGFBR2*
ENSG00000163513	GO:0018105	peptidyl-serine phosphorylation	*TGFBR2*
ENSG00000163513	GO:0018107	peptidyl-threonine phosphorylation	*TGFBR2*
ENSG00000163513	GO:0023014	signal transduction by phosphorylation	*TGFBR2*
ENSG00000163513	GO:0030324	lung development	*TGFBR2*
ENSG00000163513	GO:0030512	negative regulation of transforming growth factor beta receptor signaling pathway	*TGFBR2*
ENSG00000163513	GO:0031100	organ regeneration	*TGFBR2*
ENSG00000163513	GO:0031435	mitogen-activated protein kinase kinase kinase binding	*TGFBR2*
ENSG00000163513	GO:0032147	activation of protein kinase activity	*TGFBR2*
ENSG00000163513	GO:0034713	type I transforming growth factor beta receptor binding	*TGFBR2*
ENSG00000163513	GO:0034714	type III transforming growth factor beta receptor binding	*TGFBR2*
ENSG00000163513	GO:0035162	embryonic hemopoiesis	*TGFBR2*

ENSG00000163513	GO:0042060	wound healing	*TGFBR2*
ENSG00000163513	GO:0042127	regulation of cell proliferation	*TGFBR2*
ENSG00000163513	GO:0042493	response to drug	*TGFBR2*
ENSG00000163513	GO:0043011	myeloid dendritic cell differentiation	*TGFBR2*
ENSG00000163513	GO:0043235	receptor complex	*TGFBR2*
ENSG00000163513	GO:0043415	positive regulation of skeletal muscle tissue regeneration	*TGFBR2*
ENSG00000163513	GO:0043627	response to estrogen	*TGFBR2*
ENSG00000163513	GO:0045121	membrane raft	*TGFBR2*
ENSG00000163513	GO:0045766	positive regulation of angiogenesis	*TGFBR2*
ENSG00000163513	GO:0046332	SMAD binding	*TGFBR2*
ENSG00000163513	GO:0046872	metal ion binding	*TGFBR2*
ENSG00000163513	GO:0048545	response to steroid hormone	*TGFBR2*
ENSG00000163513	GO:0048565	digestive tract development	*TGFBR2*
ENSG00000163513	GO:0048661	positive regulation of smooth muscle cell proliferation	*TGFBR2*
ENSG00000163513	GO:0048701	embryonic cranial skeleton morphogenesis	*TGFBR2*
ENSG00000163513	GO:0050431	transforming growth factor beta binding	*TGFBR2*
ENSG00000163513	GO:0051138	positive regulation of NK T cell differentiation	*TGFBR2*
ENSG00000163513	GO:0051216	cartilage development	*TGFBR2*
ENSG00000163513	GO:0060021	palate development	*TGFBR2*
ENSG00000163513	GO:0060044	negative regulation of cardiac muscle cell proliferation	*TGFBR2*
ENSG00000163513	GO:0060389	pathway-restricted SMAD protein phosphorylation	*TGFBR2*
ENSG00000163513	GO:0060425	lung morphogenesis	*TGFBR2*
ENSG00000163513	GO:0060433	bronchus development	*TGFBR2*
ENSG00000163513	GO:0060434	bronchus morphogenesis	*TGFBR2*
ENSG00000163513	GO:0060439	trachea morphogenesis	*TGFBR2*
ENSG00000163513	GO:0060440	trachea formation	*TGFBR2*
ENSG00000163513	GO:0060443	mammary gland morphogenesis	*TGFBR2*
ENSG00000163513	GO:0060463	lung lobe morphogenesis	*TGFBR2*
ENSG00000163513	GO:0070022	transforming growth factor beta receptor homodimeric complex	*TGFBR2*
ENSG00000163513	GO:0070723	response to cholesterol	*TGFBR2*
ENSG00000163513	GO:1990086	lens fiber cell apoptotic process	*TGFBR2*
ENSG00000163513	GO:2000379	positive regulation of reactive oxygen species metabolic process	*TGFBR2*
ENSG00000069702	GO:0001568	blood vessel development	*TGFBR3*
ENSG00000069702	GO:0001666	response to hypoxia	*TGFBR3*
ENSG00000069702	GO:0001701	in utero embryonic development	*TGFBR3*
ENSG00000069702	GO:0001824	blastocyst development	*TGFBR3*
ENSG00000069702	GO:0001837	epithelial to mesenchymal transition	*TGFBR3*
ENSG00000069702	GO:0001889	liver development	*TGFBR3*
ENSG00000069702	GO:0003007	heart morphogenesis	*TGFBR3*
ENSG00000069702	GO:0005024	transforming growth factor beta-activated receptor activity	*TGFBR3*
ENSG00000069702	GO:0005114	type II transforming growth factor beta receptor binding	*TGFBR3*
ENSG00000069702	GO:0005160	transforming growth factor beta receptor binding	*TGFBR3*
ENSG00000069702	GO:0005515	protein binding	*TGFBR3*
ENSG00000069702	GO:0005539	glycosaminoglycan binding	*TGFBR3*
ENSG00000069702	GO:0005578	proteinaceous extracellular matrix	*TGFBR3*

ENSG00000069702	GO:0005615	extracellular space	*TGFBR3*
ENSG00000069702	GO:0005737	cytoplasm	*TGFBR3*
ENSG00000069702	GO:0005783	endoplasmic reticulum	*TGFBR3*
ENSG00000069702	GO:0005887	integral component of plasma membrane	*TGFBR3*
ENSG00000069702	GO:0006461	protein complex assembly	*TGFBR3*
ENSG00000069702	GO:0006955	immune response	*TGFBR3*
ENSG00000069702	GO:0007179	transforming growth factor beta receptor signaling pathway	*TGFBR3*
ENSG00000069702	GO:0007181	transforming growth factor beta receptor complex assembly	*TGFBR3*
ENSG00000069702	GO:0008201	heparin binding	*TGFBR3*
ENSG00000069702	GO:0009897	external side of plasma membrane	*TGFBR3*
ENSG00000069702	GO:0009986	cell surface	*TGFBR3*
ENSG00000069702	GO:0010633	negative regulation of epithelial cell migration	*TGFBR3*
ENSG00000069702	GO:0010719	negative regulation of epithelial to mesenchymal transition	*TGFBR3*
ENSG00000069702	GO:0015026	coreceptor activity	*TGFBR3*
ENSG00000069702	GO:0016049	cell growth	*TGFBR3*
ENSG00000069702	GO:0016477	cell migration	*TGFBR3*
ENSG00000069702	GO:0017134	fibroblast growth factor binding	*TGFBR3*
ENSG00000069702	GO:0030165	PDZ domain binding	*TGFBR3*
ENSG00000069702	GO:0030509	BMP signaling pathway	*TGFBR3*
ENSG00000069702	GO:0030511	positive regulation of transforming growth factor beta receptor signaling pathway	*TGFBR3*
ENSG00000069702	GO:0030512	negative regulation of transforming growth factor beta receptor signaling pathway	*TGFBR3*
ENSG00000069702	GO:0031100	organ regeneration	*TGFBR3*
ENSG00000069702	GO:0032354	response to follicle-stimulating hormone	*TGFBR3*
ENSG00000069702	GO:0034673	inhibin-betaglycan-ActRII complex	*TGFBR3*
ENSG00000069702	GO:0034695	response to prostaglandin E	*TGFBR3*
ENSG00000069702	GO:0034699	response to luteinizing hormone	*TGFBR3*
ENSG00000069702	GO:0035556	intracellular signal transduction	*TGFBR3*
ENSG00000069702	GO:0043393	regulation of protein binding	*TGFBR3*
ENSG00000069702	GO:0046332	SMAD binding	*TGFBR3*
ENSG00000069702	GO:0048185	activin binding	*TGFBR3*
ENSG00000069702	GO:0050431	transforming growth factor beta binding	*TGFBR3*
ENSG00000069702	GO:0050680	negative regulation of epithelial cell proliferation	*TGFBR3*
ENSG00000069702	GO:0051092	positive regulation of NF-kappaB transcription factor activity	*TGFBR3*
ENSG00000069702	GO:0051271	negative regulation of cellular component movement	*TGFBR3*
ENSG00000069702	GO:0055010	ventricular cardiac muscle tissue morphogenesis	*TGFBR3*
ENSG00000069702	GO:0060021	palate development	*TGFBR3*
ENSG00000069702	GO:0060038	cardiac muscle cell proliferation	*TGFBR3*
ENSG00000069702	GO:0060216	definitive hemopoiesis	*TGFBR3*
ENSG00000069702	GO:0060317	cardiac epithelial to mesenchymal transition	*TGFBR3*
ENSG00000069702	GO:0060318	definitive erythrocyte differentiation	*TGFBR3*
ENSG00000069702	GO:0060347	heart trabecula formation	*TGFBR3*
ENSG00000069702	GO:0060389	pathway-restricted SMAD protein	*TGFBR3*

		phosphorylation	
ENSG00000069702	GO:0070062	extracellular vesicular exosome	*TGFBR3*
ENSG00000069702	GO:0070123	transforming growth factor beta receptor activity, type III	*TGFBR3*
ENSG00000159445	GO:0005739	mitochondrion	*THEM4*
ENSG00000159445	GO:0005743	mitochondrial inner membrane	*THEM4*
ENSG00000159445	GO:0005758	mitochondrial intermembrane space	*THEM4*
ENSG00000159445	GO:0005829	cytosol	*THEM4*
ENSG00000159445	GO:0005886	plasma membrane	*THEM4*
ENSG00000159445	GO:0006631	fatty acid metabolic process	*THEM4*
ENSG00000159445	GO:0007173	epidermal growth factor receptor signaling pathway	*THEM4*
ENSG00000159445	GO:0008286	insulin receptor signaling pathway	*THEM4*
ENSG00000159445	GO:0008543	fibroblast growth factor receptor signaling pathway	*THEM4*
ENSG00000159445	GO:0016290	palmitoyl-CoA hydrolase activity	*THEM4*
ENSG00000159445	GO:0032587	ruffle membrane	*THEM4*
ENSG00000159445	GO:0038095	Fc-epsilon receptor signaling pathway	*THEM4*
ENSG00000159445	GO:0043491	protein kinase B signaling	*THEM4*
ENSG00000159445	GO:0045087	innate immune response	*THEM4*
ENSG00000159445	GO:0048011	neurotrophin TRK receptor signaling pathway	*THEM4*
ENSG00000159445	GO:0048015	phosphatidylinositol-mediated signaling	*THEM4*
ENSG00000159445	GO:1902108	regulation of mitochondrial membrane permeability involved in apoptotic process	*THEM4*
ENSG00000171914	GO:0001726	ruffle	*TLN2*
ENSG00000171914	GO:0003779	actin binding	*TLN2*
ENSG00000171914	GO:0005158	insulin receptor binding	*TLN2*
ENSG00000171914	GO:0005198	structural molecule activity	*TLN2*
ENSG00000171914	GO:0005200	structural constituent of cytoskeleton	*TLN2*
ENSG00000171914	GO:0005515	protein binding	*TLN2*
ENSG00000171914	GO:0005737	cytoplasm	*TLN2*
ENSG00000171914	GO:0005856	cytoskeleton	*TLN2*
ENSG00000171914	GO:0005886	plasma membrane	*TLN2*
ENSG00000171914	GO:0005911	cell-cell junction	*TLN2*
ENSG00000171914	GO:0005916	fascia adherens	*TLN2*
ENSG00000171914	GO:0005925	focal adhesion	*TLN2*
ENSG00000171914	GO:0007016	cytoskeletal anchoring at plasma membrane	*TLN2*
ENSG00000171914	GO:0007043	cell-cell junction assembly	*TLN2*
ENSG00000171914	GO:0007155	cell adhesion	*TLN2*
ENSG00000171914	GO:0015629	actin cytoskeleton	*TLN2*
ENSG00000171914	GO:0045202	synapse	*TLN2*
ENSG00000171914	GO:0051015	actin filament binding	*TLN2*
ENSG00000138111	GO:0016021	integral component of membrane	*TMEM180*
ENSG00000187806	GO:0016021	integral component of membrane	*TMEM202*
ENSG00000141655	GO:0001503	ossification	*TNFRSF11A*
ENSG00000141655	GO:0002250	adaptive immune response	*TNFRSF11A*
ENSG00000141655	GO:0002548	monocyte chemotaxis	*TNFRSF11A*
ENSG00000141655	GO:0004872	receptor activity	*TNFRSF11A*
ENSG00000141655	GO:0004888	transmembrane signaling receptor activity	*TNFRSF11A*
ENSG00000141655	GO:0005031	tumor necrosis factor-activated receptor activity	*TNFRSF11A*
ENSG00000141655	GO:0005515	protein binding	*TNFRSF11A*

ENSG00000141655	GO:0007165	signal transduction	*TNFRSF11A*
ENSG00000141655	GO:0007267	cell-cell signaling	*TNFRSF11A*
ENSG00000141655	GO:0007275	multicellular organismal development	*TNFRSF11A*
ENSG00000141655	GO:0008284	positive regulation of cell proliferation	*TNFRSF11A*
ENSG00000141655	GO:0009314	response to radiation	*TNFRSF11A*
ENSG00000141655	GO:0009897	external side of plasma membrane	*TNFRSF11A*
ENSG00000141655	GO:0009986	cell surface	*TNFRSF11A*
ENSG00000141655	GO:0016021	integral component of membrane	*TNFRSF11A*
ENSG00000141655	GO:0019955	cytokine binding	*TNFRSF11A*
ENSG00000141655	GO:0030316	osteoclast differentiation	*TNFRSF11A*
ENSG00000141655	GO:0032496	response to lipopolysaccharide	*TNFRSF11A*
ENSG00000141655	GO:0033209	tumor necrosis factor-mediated signaling pathway	*TNFRSF11A*
ENSG00000141655	GO:0034097	response to cytokine	*TNFRSF11A*
ENSG00000141655	GO:0034612	response to tumor necrosis factor	*TNFRSF11A*
ENSG00000141655	GO:0043123	positive regulation of I-kappaB kinase/NF-kappaB signaling	*TNFRSF11A*
ENSG00000141655	GO:0043507	positive regulation of JUN kinase activity	*TNFRSF11A*
ENSG00000141655	GO:0046872	metal ion binding	*TNFRSF11A*
ENSG00000141655	GO:0048535	lymph node development	*TNFRSF11A*
ENSG00000141655	GO:0051091	positive regulation of sequence-specific DNA binding transcription factor activity	*TNFRSF11A*
ENSG00000141655	GO:0051092	positive regulation of NF-kappaB transcription factor activity	*TNFRSF11A*
ENSG00000141655	GO:0060086	circadian temperature homeostasis	*TNFRSF11A*
ENSG00000141655	GO:0060749	mammary gland alveolus development	*TNFRSF11A*
ENSG00000141655	GO:0070555	response to interleukin-1	*TNFRSF11A*
ENSG00000141655	GO:0071812	positive regulation of fever generation by positive regulation of prostaglandin secretion	*TNFRSF11A*
ENSG00000141655	GO:0071847	TNFSF11-mediated signaling pathway	*TNFRSF11A*
ENSG00000141655	GO:0071848	positive regulation of ERK1 and ERK2 cascade via TNFSF11-mediated signaling	*TNFRSF11A*
ENSG00000198846	GO:0003677	DNA binding	*TOX*
ENSG00000198846	GO:0005634	nucleus	*TOX*
ENSG00000196511	GO:0004788	thiamine diphosphokinase activity	*TPK1*
ENSG00000196511	GO:0005524	ATP binding	*TPK1*
ENSG00000196511	GO:0005829	cytosol	*TPK1*
ENSG00000196511	GO:0006766	vitamin metabolic process	*TPK1*
ENSG00000196511	GO:0006767	water-soluble vitamin metabolic process	*TPK1*
ENSG00000196511	GO:0006772	thiamine metabolic process	*TPK1*
ENSG00000196511	GO:0009229	thiamine diphosphate biosynthetic process	*TPK1*
ENSG00000196511	GO:0016301	kinase activity	*TPK1*
ENSG00000196511	GO:0016310	phosphorylation	*TPK1*
ENSG00000196511	GO:0030975	thiamine binding	*TPK1*
ENSG00000196511	GO:0042723	thiamine-containing compound metabolic process	*TPK1*
ENSG00000196511	GO:0044281	small molecule metabolic process	*TPK1*
ENSG00000072657	GO:0004177	aminopeptidase activity	*TRHDE*
ENSG00000072657	GO:0005887	integral component of plasma membrane	*TRHDE*

ENSG00000072657	GO:0006508	proteolysis	*TRHDE*
ENSG00000072657	GO:0007165	signal transduction	*TRHDE*
ENSG00000072657	GO:0007267	cell-cell signaling	*TRHDE*
ENSG00000072657	GO:0008237	metallopeptidase activity	*TRHDE*
ENSG00000072657	GO:0008270	zinc ion binding	*TRHDE*
ENSG00000072657	GO:0070062	extracellular vesicular exosome	*TRHDE*
ENSG00000032389	GO:0005515	protein binding	*TSSC1*
ENSG00000170035	GO:0004842	ubiquitin-protein transferase activity	*UBE2E3*
ENSG00000170035	GO:0005524	ATP binding	*UBE2E3*
ENSG00000170035	GO:0005634	nucleus	*UBE2E3*
ENSG00000170035	GO:0005737	cytoplasm	*UBE2E3*
ENSG00000170035	GO:0008152	metabolic process	*UBE2E3*
ENSG00000170035	GO:0016881	acid-amino acid ligase activity	*UBE2E3*
ENSG00000170035	GO:0040008	regulation of growth	*UBE2E3*
ENSG00000170035	GO:0070534	protein K63-linked ubiquitination	*UBE2E3*
ENSG00000170035	GO:0070936	protein K48-linked ubiquitination	*UBE2E3*
ENSG00000170035	GO:0070979	protein K11-linked ubiquitination	*UBE2E3*
ENSG00000140455	GO:0000122	negative regulation of transcription from RNA polymerase II promoter	*USP3*
ENSG00000140455	GO:0000278	mitotic cell cycle	*USP3*
ENSG00000140455	GO:0000790	nuclear chromatin	*USP3*
ENSG00000140455	GO:0001046	core promoter sequence-specific DNA binding	*USP3*
ENSG00000140455	GO:0003682	chromatin binding	*USP3*
ENSG00000140455	GO:0004221	ubiquitin thiolesterase activity	*USP3*
ENSG00000140455	GO:0004843	ubiquitin-specific protease activity	*USP3*
ENSG00000140455	GO:0005634	nucleus	*USP3*
ENSG00000140455	GO:0006281	DNA repair	*USP3*
ENSG00000140455	GO:0006511	ubiquitin-dependent protein catabolic process	*USP3*
ENSG00000140455	GO:0008234	cysteine-type peptidase activity	*USP3*
ENSG00000140455	GO:0008270	zinc ion binding	*USP3*
ENSG00000140455	GO:0016578	histone deubiquitination	*USP3*
ENSG00000140455	GO:0031647	regulation of protein stability	*USP3*
ENSG00000140455	GO:0036459	ubiquitinyl hydrolase activity	*USP3*
ENSG00000140455	GO:0042393	histone binding	*USP3*
ENSG00000118369	GO:0004843	ubiquitin-specific protease activity	*USP35*
ENSG00000118369	GO:0005488	binding	*USP35*
ENSG00000118369	GO:0006511	ubiquitin-dependent protein catabolic process	*USP35*
ENSG00000118369	GO:0008234	cysteine-type peptidase activity	*USP35*
ENSG00000118369	GO:0016579	protein deubiquitination	*USP35*
ENSG00000118369	GO:0036459	ubiquitinyl hydrolase activity	*USP35*
ENSG00000117533	GO:0000139	Golgi membrane	*VAMP4*
ENSG00000117533	GO:0005764	lysosome	*VAMP4*
ENSG00000117533	GO:0005768	endosome	*VAMP4*
ENSG00000117533	GO:0005794	Golgi apparatus	*VAMP4*
ENSG00000117533	GO:0005802	trans-Golgi network	*VAMP4*
ENSG00000117533	GO:0009986	cell surface	*VAMP4*
ENSG00000117533	GO:0016021	integral component of membrane	*VAMP4*
ENSG00000117533	GO:0016192	vesicle-mediated transport	*VAMP4*
ENSG00000117533	GO:0031201	SNARE complex	*VAMP4*
ENSG00000117533	GO:0035493	SNARE complex assembly	*VAMP4*
ENSG00000117533	GO:0042996	regulation of Golgi to plasma membrane protein transport	*VAMP4*
ENSG00000117533	GO:0043001	Golgi to plasma membrane protein	*VAMP4*

		transport	
ENSG00000075035	GO:0016021	integral component of membrane	*WSCD2*
ENSG00000206579	GO:0016021	integral component of membrane	*XKR4*
ENSG00000145817	GO:0005789	endoplasmic reticulum membrane	*YIPF5*
ENSG00000145817	GO:0015031	protein transport	*YIPF5*
ENSG00000145817	GO:0016020	membrane	*YIPF5*
ENSG00000145817	GO:0016021	integral component of membrane	*YIPF5*
ENSG00000145817	GO:0016192	vesicle-mediated transport	*YIPF5*
ENSG00000145817	GO:0030134	ER to Golgi transport vesicle	*YIPF5*
ENSG00000145817	GO:0032580	Golgi cisterna membrane	*YIPF5*
ENSG00000145817	GO:0042175	nuclear outer membrane-endoplasmic reticulum membrane network	*YIPF5*
ENSG00000145817	GO:0060628	regulation of ER to Golgi vesicle-mediated transport	*YIPF5*
ENSG00000145817	GO:0070971	endoplasmic reticulum exit site	*YIPF5*
ENSG00000181722	GO:0000122	negative regulation of transcription from RNA polymerase II promoter	*ZBTB20*
ENSG00000181722	GO:0001078	RNA polymerase II core promoter proximal region sequence-specific DNA binding transcription factor activity involved in negative regulation of transcription	*ZBTB20*
ENSG00000181722	GO:0003677	DNA binding	*ZBTB20*
ENSG00000181722	GO:0005515	protein binding	*ZBTB20*
ENSG00000181722	GO:0005634	nucleus	*ZBTB20*
ENSG00000181722	GO:0006351	transcription, DNA-templated	*ZBTB20*
ENSG00000181722	GO:0046872	metal ion binding	*ZBTB20*