

## **Final Technical Progress Report:**

### **i. Summary of project aims: The aim of this proposal is to identify genetic associations, and ultimately identify risk alleles in inflammatory bowel disease (IBD) patients on chromosomes 3q27-28.**

The largest peak overall in our large North American IBD linkage study was in the chromosome 3q27-28 region, a region demonstrating suggestive evidence for linkage in our initial linkage analysis. Importantly, the linkage evidence on chromosome 3q27-28 is observed most significantly in Jewish IBD families, and represents the most significant region of linkage evidence in Jewish cohorts. We also present data from a meta-analysis combining linkage data from 1253 IBD affected relative pairs implicating the chromosome 3q region. Because the Ashkenazi Jewish population has a significantly higher prevalence of IBD and demonstrates population substructure compared to the non-Jewish European ancestry cohort, we hypothesize that significant IBD susceptibility gene(s) reside in this region. We further hypothesize that risk alleles and linkage disequilibrium patterns may be distinct between Jewish and non-Jewish European ancestry cohorts in this region of linkage.

### **ii. Accomplishments towards meeting these aims**

We completed an initial fine mapping effort over a 5 Mb region (186.6 Mb to 191.7 Mb, gene-based) based on existing HapMap Phase I selected SNPs (single nucleotide polymorphisms) in Jewish and non-Jewish IBD cohorts. We observed nominal evidence for association in a few candidate genes as listed below, with marked differences in allele frequencies and evidence for association observed in Jewish vs. non-Jewish cohorts, typically with greater evidence for association observed in Jewish cohorts. Importantly, linkage disequilibrium in Jewish cohorts was significantly greater compared to the non-Jewish group. We hypothesize that susceptibility alleles in multiple genes in this region contribute to the linkage signal in this region.

Follow-up studies will utilize the additional SNPs available from the Perlegen and HapMap Phase II SNPs. In addition, a key effort will involve typing the following regions in independent Jewish IBD cohorts, as most of the evidence for association in this region is observed in this group. Finally, deep resequencing efforts in this region will be performed to identify rare (minor allele frequencies less than 5%) susceptibility alleles.

Ethnic differences in IBD risk alleles may prove to be the rule rather than the exception for complex, multigenic disorders. The Arg702Trp variant in the NOD2 (CARD15) gene does not demonstrate significant evidence for association in Jewish CD. The other, well-replicated IBD association, IBD5 (on chromosome 5q) similarly does not demonstrate any evidence for association of the risk haplotype in this region. Coding region polymorphisms in pattern recognition receptors mediating host response to intestinal bacteria are of particular importance in defining the pathogenesis of IBD. A stop codon in the TLR5 receptor has been identified which results in cellular non-responsiveness to flagellin. It has recently been reported that flagellin is an immunodominant epitope in Crohn's disease (CD), but not ulcerative colitis (UC). We demonstrate that the TLR5-stop carriage mediates non-responsiveness to flagellin and is observed in a lower frequency in Jewish compared to non-Jewish populations. Furthermore, in Jewish CD, carriage of TLR5-stop protects against the development of CD. This demonstrates that natural acquisition of immune responses to flagellin are regulated by TLR5 and suggest that immune responses to flagellin are not merely associated with CD but rather increase one's potential for developing CD.

### **iii. A list of significant results**

**a) Throughout 186.6 Mb to 191.7 Mb (chr 3q27-8) region on chromosome 3q, significant ethnic differences are observed between non-Jewish and Jewish CD and UC.**

Fine-mapping in a family-based association cohort (i.e. affected child/sib pair plus both parents typed) was performed and marked differences in allelic frequencies were observed throughout the 5 Mb region. Of particular interest were the consistently different allele frequencies observed throughout the SIAT1 gene, with some differences observed for the MASP-1 gene region.

marker	Jewish CD	Jewish UC	non-Jewish CD	non-Jewish UC	P-value
rs3864106-SIAT1	0.31	0.22	0.42	0.46	0
rs4686820-SIAT1	0.23	0.24	0.11	0.13	0
rs11712248-SIAT1	0.12	0.11	0.22	0.21	6.00E-04
rs9828559-SIAT1	0.98	0.96	0.9	0.93	8.00E-04
rs6444192-SIAT1	0.34	0.26	0.45	0.53	0
rs17718970-SIAT1	0.12	0.23	0.26	0.39	0
rs16848727-SIAT1	0.33	0.45	0.22	0.24	1.00E-04
rs7619989-SIAT1	0.28	0.31	0.4	0.43	4.00E-04
rs4686837-SIAT1	0.63	0.61	0.76	0.75	1.00E-04
rs9941987-SIAT1	0.05	0.09	0.18	0.13	0
rs12638323-SIAT1	0.23	0.29	0.45	0.38	0
rs9821439-SIAT1	0.22	0.26	0.44	0.38	0
rs9716-SIAT1	0.12	0.11	0.27	0.24	0
rs4686849-SIAT1	0.23	0.29	0.42	0.39	0
rs2008826	0.21	0.26	0.4	0.38	0
rs710459-MASP1	0.3	0.41	0.43	0.46	9.00E-04

**b) Through the 3q27-8 region, we observe significantly greater evidence for linkage disequilibrium in Jewish compared to non-Jewish cohorts**



We observe greater linkage disequilibrium (LD) in Jewish compared to non-Jewish IBD cohorts. In this LD plot, adjacent markers are plotted along the diagonal. The squares occurring at the junction of two points along the diagonal reflects the extent of LD for that pair of markers, with darker squares corresponding to higher degrees of LD. Note that for closely related markers (that is, close to each other along the diagonal), there is significant LD, reflected by

the preponderance of dark squares close to the diagonal. However, LD can extend over longer distances, as reflected by the dark squares in the *interior* of the triangle. Note that there is significantly greater LD in the interior of the triangle in the Jewish cohort compared to the non-Jewish group. We hypothesize that multiple risk alleles in adjacent genes contribute to the linkage signal that we observed in this region.

**c. Association analysis of gene-based SNPs demonstrates modest evidence for association in IMP-2 and LPP genes**

Of interest in this region is the MASP-1 (mannan-binding lectin serine peptidase 1), a complement dependent bactericidal factor and results from a deep resequencing effort in this region will be presented. We demonstrate

Marker	Gene symbol	location	IBD p-value	CD p-value
rs2290066	IMP-2	186,766,755	0.009	0.06
rs1053980	LPP	189,648,374	0.02	0.0038
rs3796293	IL-1RAP	191,653,006	0.16	0.02

nominal evidence for association for the IMP-2 (insulin-like growth factor 2 mRNA binding protein), LPP (LIM domain containing preferred translocation partner in lipoma), and IL-1RAP (interleukin receptor accessory protein) genes. Of note is that despite markedly different allele frequencies throughout the SIAT1 gene, we thus far have not identified significant IBD association in this region. Studies using independent cohorts of Jewish IBD patients are ongoing.

**d. The TLR5-stop codon reduces acquired immune responses to bacterial flagellin and, in Jewish cohorts, protects against the development of CD.**

In healthy subjects, persons carrying TLR5-stop had significantly lower levels of flagellin-specific IgG and IgA but had similar levels of total and LPS-specific Ig. Moreover, we observed that among Jewish subjects, the carriage rate of TLR5-stop (in heterozygous state) was significantly less in CD patients, but not ulcerative colitis (UC) patients, compared to unaffected relatives and unrelated controls (5.4%, 0.9%, 6.0%, and 6.5% for unaffected relatives, CD, UC, and unrelated Jewish controls respectively, n= 296, 215, 185, and 416; p=0.037 by likelihood calculation for CD vs. controls) indicating that TLR5-stop can protect persons of Jewish ethnicity against CD. We did not observe a significant association of TLR5-stop with CD in a non-Jewish cohort (11.1%, 10.4%, and 11.7% for unaffected relatives, CD, and UC; n=841, 543, and 300 for unaffected relatives).

## **vi. Lay summary of progress**

The work funded by the BMRP allowed us to further define genetic variation in the chromosome 3q region and define ethnic differences with respect genetic associations to inflammatory bowel disease (IBD). We focused on this region because in our very large study, we observed the greatest degree of sharing in this region between relatives having IBD. In particular, this region demonstrated a great amount of sharing in Jewish affected relative pairs.

We observed marked differences in the frequencies of genetic markers in this region, and observed greater degrees of chromosomal sharing between affected Jewish individuals compared to non-Jewish European ancestry individuals. Therefore, we hypothesize that the susceptibility alleles in this region will be distinct between Jewish and non-Jewish IBD, and that multiple genes contribute to the linkage we observed in this region. We observe some evidence for association to genes involved in growth, differentiation and cytokine regulation and response, and confirmatory studies are ongoing.

Unexpectedly, we now believe that susceptibility and protective alleles will have different patterns of association in different ethnic groups. For example we have shown that an allele which truncates a protein receptor mediating responses to a bacterial component actually protects against the development of Crohn's disease in Jewish populations. We know that, while intestinal bacteria are required for normal gastrointestinal development, in IBD, the inflammatory response is largely driven by an excessive immune response to these normal bacteria. We anticipate that the complete genetic elucidation of susceptibility alleles will involve an understanding of the unique combination of factors in individual patients. These different contributing genetic factors will, at some point, translate into individualized therapeutic approaches.