(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
2 October 2008 (02.10.2008)

PCT

(10) International Publication Number
WO 2008/118969 A2

(51) International Patent Classification:
C12Q 1/68 (2006.01)

(21) International Application Number:
PCT/US2008/058281

(22) International Filing Date: 26 March 2008 (26.03.2008)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/919,953 26 March 2007 (26.03.2007) US
60/907,543 6 April 2007 (06.04.2007) US

(71) Applicant *(for all designated States except US)*: **GENIZON BIOSCIENCES INC** [CA/CA]; 880 McCaffrey Street, Ville St-Laurent, Québec H4T 2C7 (CA).

(72) Inventors; and

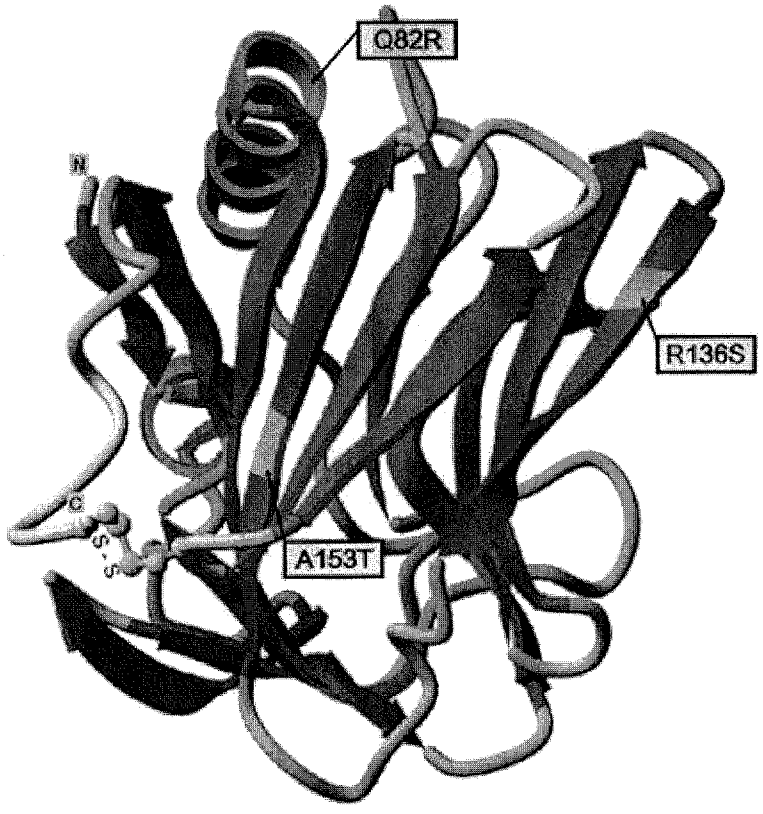
(75) Inventors/Applicants *(for US only)*: **SCHREIBER, Stefan** [DE/CA]; c/o GENIZON BIOSCIENCES INC, 880 McCaffrey Street, Ville St-Laurent, Québec H4T 2C7 (CA). **ROSENSTIEL, Philip** [DE/CA]; c/o GENIZON BIOSCIENCES INC, 880 McCaffrey Street, Ville St-Laurent, Québec H4T 2C7 (CA). **RUTHER, Andreas** [DE/CA]; c/o GENIZON BIOSCIENCES INC, 880 McCaffrey Street, Ville St-Laurent, Québec H4T 2C7 (CA). **FRANKE, Andre** [DE/CA]; c/o GENIZON BIOSCIENCES INC, 880 McCaffrey Street, Ville St-Laurent, Québec H4T 2C7 (CA). **RAELSON, John** [US/CA]; c/o GENIZON BIOSCIENCES INC, 880 McCaffrey Street, Ville St-Laurent, Québec H4T 2C7 (CA). **LITTLE, Randall D.** [US/CA]; c/o GENIZON BIOSCIENCES INC, 880 McCaffrey Street, Ville St-Laurent, Québec H4T 2C7 (CA). **KEITH, Tim** [US/US]; c/o GENIZON BIOSCIENCES INC, 880 McCaffrey Street, Ville St-Laurent, Québec H4T 2C7 (CA). **FOURNIER, Helene** [CA/CA]; c/o GENIZON BIOSCIENCES INC, 880 McCaffrey Street, Ville St-Laurent, Québec H4T 2C7 (CA).

(74) Agent: **ROCHON, Marie-Hélène**; OGILVY RENAULT LLP/S.E.N.C.R.L., s.r.l., 1981 McGill College, Suite 1600, Montreal, Québec H3A 2Y3 (CA).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM,

[Continued on next page]

(54) Title: METHODS AND AGENTS FOR EVALUATING INFLAMMATORY BOWEL DISEASE, AND TARGETS FOR TREATMENT



(57) Abstract: The invention provides methods for evaluating irritable bowel disease (IBD), including Crohn Disease and Ulcerative Colitis, methods for determining a patient's susceptibility to developing an IBD, and methods for determining a patient's IBD genotype. The invention includes methods, polynucleotides, polypeptides, and antibodies relating to disclosed variants of, and polymorphisms in, the nel-like 1 precursor (NELL1), as well as the 5p13.1 locus, and other genes disclosed herein to be associated with IBD. Thus, the invention provides diagnostic and/or therapeutic targets for IBD, as well as diagnostic and therapeutic agents for IBD.

WO 2008/118969 A2

AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MT, NL, NO, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— *without international search report and to be republished upon receipt of that report*

— *with sequence listing part of description published separately in electronic form and available upon request from the International Bureau*

METHODS AND AGENTS FOR EVALUATING INFLAMMATORY BOWEL DISEASE, AND TARGETS FOR TREATMENT

PRIORITY

[01] This application claims priority to US Provisional Application Nos. 60/919,953 filed March 26, 2007, and 60/907,543 filed April 6, 2007, each of which is hereby incorporated by reference in its entirety.

BACKGROUND

[02] An estimated 1.4 million individuals in the United States and 2.2 million individuals in Europe suffer from inflammatory bowel disease [1], a life-long disease that occurs in the form of one of two major sub-phenotypes, Crohn disease (CD) or ulcerative colitis (UC). The pathophysiology of IBD is characterized by a highly activated state of the mucosal immune system and excessive mucosal destruction. The enteric flora appears to play a key role as a stimulating agent [2]. Familial clustering [3,4] and an increased concordance rate of IBD among monozygotic twins [5,6] are hallmarks of the genetic aetiology of IBD, a notion that is further supported by the discovery of several disease genes. These include *NOD2* [7–9] (*IBD1*), a risk haplotype in the *5q31* (*IBD5*) locus [10,11], *DLG5* [12,13], *TNFSF15* [14], *ATG16L1* [15], *CARD4* [16], and *IL23R* [17].

[03] Yamazaki and colleagues reported the first genome-wide association scan (GWS) for CD, which resulted in the identification of associated polymorphisms in the *TNFSF15* gene [14]. Two other GWSs reported the novel CD susceptibility loci *IL23R* [17] and *5p13.1* [18]. A genome-wide candidate gene analysis was performed using 19,779 non-synonymous SNPs, which led to the identification of a common variant (T300A) in the *ATG16L1* gene as predisposing to CD [15], a finding that was later replicated by four other groups [18–21].

[04] Although studies have identified some CD-associated genetic variants, these susceptibility loci explain only a fraction of the heritability of the disease. Thus, identification of additional risk loci for IBD are needed to provide effective diagnostic, prognostic, and therapeutic targets for IBD, including Crohn Disease and ulcerative colitis.

SUMMARY OF THE INVENTION

[05] Disclosed are risk loci for IBD, which were identified through a multi-stage genome-wide association scan in 393 German cases and 399 German population-representative controls, using the Affymetrix GeneChip® Human Mapping 100K Set [22]. In order to enrich the samples with risk alleles [23] and to reduce phenotypic heterogeneity, CD patients in the

GWS were selected for a ''severe'' phenotype, including a positive IBD family history, age of onset ≤25 years, and no change in diagnosis over the prior five years. The SNPs representing the top 200 association leads were re-genotyped in both a large independent German case control sample and a family-based sample comprising 375 nuclear families. In addition to replicating *NOD2*, *IBD5*, and *5p13.1*, a novel susceptibility locus was identified on chromosome *11p15.1*, namely the nel-like 1 precursor-encoding gene (*NELL1*).

[06] Thus, in one aspect, the present invention provides a method for evaluating irritable bowel disease (IBD) in a patient suspected of having an IBD, including Crohn Disease. This aspect further relates to a method of determining a patient's susceptibility to developing an IBD, and to methods for determining a patient's IBD genotype. In accordance with some embodiments, the method comprises determining the presence or absence, in a patient's biological sample, of at least one mutation associated with IBD in each of at least two genes listed in Table 1. In accordance with some embodiments, the method involves determining the presence or absence of at least one single nucleotide polymorphism (SNP) in a biological sample from the patient, including at least one the SNP listed in Tables 1-5.

[07] In a related aspect for evaluating IBD, the invention involves determining the presence or absence of at least one mutation in the *11p15.1* locus that is associated with IBD, including mutations in the gene encoding the nel-like 1 precursor (*NELL1*), in a biological sample from the patient. Such mutations include SNPs localized to the NELL1 gene as disclosed in Tables 1, 2, and 4, as well as mutations encoding the variants R136S, A153T, and/or R354W of the NELL1 polypeptide.

[08] In another related aspect for evaluating IBD, the invention involves determining the presence or absence in a patient's biological sample of: at least one mutation in the *5p13.1* locus that is associated with IBD, including *PTGER4* (upstream), and including SNPs disclosed in Tables 1, 3, and 5; and/or at least one mutation associated with IBD in *ITGB6* (upstream); and/or at least one mutation associated with IBD in *GRM8* (downstream); and/or at least one mutation associated with IBD in *OR5V1* (downstream), at least one mutation associated with IBD in *PPP3R2* (downstream); and/or at least one mutation associated with IBD in *NM_152575* (upstream); and/or at least one mutation associated with IBD in *HNF4G* (intron).

[09] In a second aspect, the invention provides novel variants of the NELL1 protein and encoding polynucleotides, vectors, and host cells. This aspect further provides antibodies recognizing, in a specific fashion, the novel NELL1 variants. Such products have use as diagnostic, prognostic, and therapeutic targets, and use as diagnostic, prognostic, and therapeutic agents for IBD, including Crohn Disease.

[010] In a third aspect, the invention provides a kit or array containing nucleic acid primers and/or probes for determining the presence and/or absence of an IBD risk genotype in a patient sample. The set of probes and/or primers may consist essentially of primers and/or probes related to evaluating an IBD genotype, and primers and/or probes related to necessary or meaningful assay controls. The kit for evaluating an IBD genotype may comprise nucleic acid probes and/or primers designed to detect ten or more SNPs associated with IBD, such as associated SNPs found in the genes listed in Table 1, and including the SNPs listed in Tables 1-5. Alternatively, the kit for evaluating an IBD genotype may contain probes and/or primers for detecting at least one mutation in *NELL1*, and optionally at least one mutation associated with an IBD in one or more of *NOD2*, the *5q31* locus, *DLG5*, *TNFSF15*, *ATG16L1*, *CARD4*, and *IL23R*. In accordance with this aspect, the kit may be a companion diagnostic kit for evaluating IBD or determining an IBD genotype in a patient, and for selecting or predicting appropriate therapeutic intervention.

DESCRIPTION OF THE DRAWINGS

[011] **Figure 1**. *Expression and localization of NELL1*. **(A)** Transcript levels of NELL1 in a set of different tissues were quantified by RT-PCR. Parallel amplification of β-actin (ACTB) is shown. Expression and localization of the NELL1 protein in healthy colonic tissue is demonstrated in sections **(B)** (20x) and **(C)** (40x; bar = 10 µm) by immunohistochemistry. Immunoreactivity is confined to mononuclear/lymphocytic cells in the lamina propria (brown DAB reaction product, arrows). A control without the primary antibody is shown in section **(D)**. No major expression differences between colonic specimen from normal controls (N) and Crohn disease (CD) were detected in the Western blot (**E**) with the same antibody as applied in sections B and C. The single detected band (90 kDa) corresponds to the predicted molecular weight of the isoform encoded by GenBank AK127805 (UniProt accession number Q92832).

[012] **Figure 2**. **(A)** Structure-based multiple sequence alignment of the N-terminal domains of NELL1 and NELL2 homologs (SEQ ID NOS: 1-8) and the N-terminal domain of human thrombospondin-1 (TSP-1) (SEQ ID NO: 9). The DSSP secondary structure assignment of the TSPN structure (PDB code 1z78, chain A) is depicted at the top of the alignment. **(B)** Multiple sequence alignment of NELL1 and NELL2 homologs (SEQ ID NOS: 1-8). Domain locations are represented at the top of the alignment (VWC domain; EGF-like domain). Alignment columns with more than 70% physicochemically similar amino acids are highlighted. Text labels point to the N-terminal signal peptide and the sequence variants.

Residue numbering in the alignment is based on complete protein sequences as derived from corresponding UniProt entries.

[013] **Figure 3**. *In silico protein analysis*. Domain architectures of NELL1/ NELL2 and TSP1. The positions of variant amino acids are annotated. Abbreviations are as follows: SP: signal peptide; TSPN: thrombospondin N-terminal domain; CC: coiled-coil region; VWC: von Willebrand factor, type C domain; EGF: EGF domain; TSP-3: thrombospondin-3 repeat; TSP_C: thrombospondin C-terminal domain.

[014] **Figure 4**. *In silico protein analysis*. Computationally derived 3D structure model of the N-terminal domain of the NELL1 protein. The model was created using the TSPN (PDB code 1z78, chain A) as a structure template for NELL1. The locations of variant amino acids as well as of two cysteines forming a disulfide bridge are annotated.

DESCRIPTION OF THE TABLES

[015] **Table 1**: Top 200 CD-associated SNPs, ranked with respect to p-values obtained in an allele- (pCCA) or genotype-based (pCCG) case-control comparison in panel A. Also included are pCCA, pCCG, and the transmission disequilibrium test results (pTDT) for the replication panel B. Nucleotide positions refer to NCBI build 34. Markers with $p \leq 0.05$ in either the case-control analysis or the transmission disequilibrium test (TDT) in replication panel B are highlighted in bold italics. SNPs with a significant result in both panel B tests are additionally marked by grey shading.

[016] **Table 2**: Fine mapping of the CD association signal at the NELL1 locus in replication panel B. The p-values of the allele-based (pCCA) and genotype-based (pCCG) association analyses of the tagging SNPs are shown, pTDT is the p-value for the transmission disequilibrium test (TDT). Lead SNPs from the initial screening (see Table 1) are highlighted by grey shading, nonsynonymous SNPs are highlighted. Polymorphisms that are significant in either the TDT or the case-control analyses, are highlighted in bold italics. Pairwise LD is listed using the metric r2 as calculated with Haploview39 and minor allele frequencies (MAF) are listed for control individuals.

[017] **Table 3**: Fine mapping of the CD association signal at the 5p13.1 locus in replication panel B. The highlighting and the column headers are the same as described in Table 2.

[018] **Table 4**: Summary of the fine mapping of the NELL1 gene locus in replication panel D. In addition to the p-value for the TDT (pTDT), the corresponding transmitted/untransmitted ratio (T:U) is listed. Other column headers plus the highlighting are the same as described in Table 2.

[019] **Table 5**: 5p13.1 fine mapping in replication panel D. For a description of the column headers and the highlighting see Table 3.

DETAILED DESCRIPTION OF THE INVENTION

[020] Crohn disease (CD), a sub-entity of inflammatory bowel disease (IBD), is a complex polygenic disorder. Although recent studies have successfully identified CD-associated genetic variants, these susceptibility loci explain only a fraction of the heritability of the disease. Disclosed is a multi-stage genome-wide scan of 393 German CD cases and 399 controls. Among the 116,161 single-nucleotide polymorphisms tested, an association with the known CD susceptibility gene *NOD2*, the *5q31* haplotype, and the recently reported CD locus at *5p13.1* were confirmed. In addition, SNP rs1793004 in the gene encoding nel-like 1 precursor (*NELL1*, chromosome *11p15.1*) showed a consistent disease-association in independent German population- and family-based samples (942 cases, 1082 controls, 375 trios). Subsequent fine mapping and replication in an independent sample of 454 French/Canadian CD trios supported the authenticity of the *NELL1* association. Further confirmation in a large German ulcerative colitis (UC) sample indicated that *NELL1* is a ubiquitous IBD susceptibility locus (combined $p<10^{-6}$; OR = 1.66, 95% CI: 1.30–2.11). The novel *5p13.1* locus was also replicated in the French/Canadian sample and in an independent UK CD patient panel (453 cases, 521 controls, combined $p<10^{-6}$ for SNP rs1992660). Several associations were replicated in at least one independent sample, point to an involvement of *ITGB6* (upstream), *GRM8* (downstream), *OR5V1* (downstream), *PPP3R2* (downstream), *NM_152575* (upstream) and *HNF4G* (intron).

Methods For Evaluating IBD And Determining A Patient's IBD Genotype

[021] In one aspect, the present invention provides a method for evaluating irritable bowel disease (IBD) in a patient suspected of having an IBD, including Crohn Disease and ulcerative colitis. This aspect further relates to a method of determining a patient's susceptibility to developing an IBD, and to methods for determining a patient's IBD genotype. In accordance with some embodiments, the method comprises determining the presence or absence in a patient's biological sample, of at least one mutation associated with IBD in each of at least two genes listed in Table 1. In accordance with some embodiments, the method involves determining the presence or absence of at least one single nucleotide polymorphism (SNP) listed in Tables 1-5 in a biological sample from the patient. The SNPs listed in Tables 1-5 are publicly available, and the corresponding nucleotide sequences are hereby incorporated by reference.

[022] In certain embodiments, the method involves determining the presence or absence of two or more SNPs selected from rs2076756 (#1 in Table 1), rs1992662 (#70 in Table 1), rs1992660 (#75 in Table 1), rs1793004 (#83 in Table 1), rs10521209 (#159 in Table 1), and rs2631372 (#163 in Table 1) in the patient sample. Such SNPs localize to the IBD-associated genes and loci of *NOD2, 5q31, 5p13.1*, and *NELL1.*

[023] With respect to SNP rs951199 in NELL1 (listed in Table 2), this SNP is further associated with sarciodosis. Thus, the present invention, as it relates to SNP rs951199, further provides a method of diagnosing sarcoidosis in a subject suspected of having the disorder or condition.

[024] In a related aspect, the invention provides a method for evaluating irritable bowel disease (IBD) in a patient suspected of having an IBD, including a method of determining a patient's susceptibility to developing an IBD, and a method for determining an IBD genotype, by using the NELL1 gene as a diagnostic or prognostic target. According to these embodiments, the invention comprises determining the presence or absence of a mutation in the gene encoding the nel-like 1 precursor (NELL1) in a biological sample from the patient, or a mutation in the corresponding 11p15.1 locus. The mutation(s) may include one or more SNPs listed in Tables 1, 2, and 4 that localize to the *NELL1* gene, and which are associated with IBD, including: the SNPs rs1793004 and rs951199, and/or the SNPs rs8176785, rs8176786, rs10500885, rs1158547, and rs1945404, and/or may include mutations encoding the Q82R, R136S, A153T, and/or R354W variants of NELL1.

[025] The *neural epidermal growth-factor-like* (*nel*) gene was first detected in neural tissue from an embryonic chicken cDNA library, and its human orthologue *NELL1* was later discovered in B-cells [48–50]. The arrangement of the functional domains of the 810 aa protein bears resemblance to thrombospondin-1 (TSP-1) and consists of a thrombospondin N-terminal domain (TSPN) and several von Willebrand factor, type C (VWC), and epidermal growth-factor (EGF) domains [51]. As NELL1 binds to, and is phosphorylated by, PKC-β1 via the EGF domains [52], it has been suggested that this protein belongs to a novel class of cell-signalling ligand molecules critical for growth and development. Re-sequencing and fine mapping revealed several non-synonymous SNPs of which the known Q82R variant and the novel R136S and A153T variants affect the TSPN domain, while R354W is located in a VWC domain (Figure 3) [51]. A153T is close to two highly conserved C-terminal cysteines forming a disulfide bond in the TSPN domain structure of TSP-1 [53] and may cause local conformational changes due to its buried position in the molecule. Generally, the TSPN domain has been shown to serve as a protein-protein interaction module, which binds membrane proteins and proteoglycans and exhibits versatile cell-specific effects on adhesion, migration, and proliferation [54,55]. Since VWC domains occur in numerous

proteins of diverse functions and are generally assumed to be involved in protein oligomerization [56], R354W may interfere with *NELL1* trimerization [51].

[026] Bone development is severely disturbed in transgenic mice, where over-expression of *NELL1* leads to craniosynostis [57] and *NELL1* deficiency manifests in skeletal defects due to reduced chondro- and osteogenesis [58]. Interestingly, osteopenia and osteoporosis are leading co-morbidities in IBD patients, even without the use of glucocorticoids [59–61].

[027] In another related aspect, the invention provides a method for evaluating irritable bowel disease (IBD) in a patient suspected of having an IBD, including a method of determining a patient's susceptibility to developing an IBD, and a method for determining an IBD genotype, by using additional loci as diagnostic or prognostic targets. According to these embodiments, the invention comprises determining the presence or absence of a mutation associated with IBD, in a patient's biological sample, in one or more of: the *5p13.1* locus (including associated SNPs of the *5p13.1* locus listed in Tables 1, 3, and 5); *PTGER4* (upstream); *ITGB6* (upstream); *GRM8* (downstream); *OR5V1* (downstream); *PPP3R2* (downstream); *NM_152575* (upstream); and/or *HNF4G* (intron). In these embodiments, the method may comprise determining the presence or absence of one or more SNPs selected from rs1992662, rs1992660, rs1553575, rs7725523, rs2925757, rs6947579, rs10487428, rs10484545, rs4743484, rs7868736, rs830772, rs272867 in a biological sample (see Table 1).

[028] In certain embodiments, the method further comprises determining the presence or absence of one or more of the following: a mutation in the CARD15 gene associated with an IBD, a mutation in the DLG5 gene associated with an IBD, a mutation in the TNFSF15 gene associated with an IBD, a mutation in the IL23R gene associated with an IBD, and/or a T300A mutation in the ATG16L1 gene. Such genes have previously been shown to be associated with an IBD, and thus, these embodiments may provide diagnostic and prognostic value in combination with those disclosed above (e.g., genes and SNPs listed in Tables 1-5). In accordance with these embodiments, the method may comprise determining the presence and/or absence of the following SNPs in a biological sample: rs2066844 (*NOD2*), rs2066845 (*NOD2*), rs2066847 (*NOD2*), rs1248696 (*DLG5*), rs2289310 (*DLG5*), rs11209026 (*IL23R*), and rs2241880 (*ATG16L1*).

[029] In accordance with the methods of evaluating IBD, the patient suspected of having an IBD may be female, or may be male, and may be of any age. In some embodiments, however, the patient has a "severe" phenotype, marked in part by onset of the disease before the age of 25, optionally with symptoms of IBD beginning at least three or at least five years prior to testing. Preferably, the patient has a family history of IBD, and/or is already

suffering from symptoms of an IBD including Crohn disease or ulcerative colitis at the time of testing. Thus, the methods of the invention, in addition to determining a likelihood of developing an IBD, aid in confirming a diagnosis of an IBD, and provide a means for determining a particular IBD genotype. Knowledge of the particular IBD genotype of the patient will aid in evaluating the likely or potential disease progression as well as the selection of an appropriate therapeutic intervention.

[030] For determining the presence or absence of mutations and/or SNPs in accordance with the invention, samples may be obtained from any part(s) of the patient's body including, but not limited to, hair, mouth, rectum, colon, scalp, blood, dermis, epidermis, and skin cells.

[031] In accordance with these aspects, the presence and/or absence of SNPs or mutations may be determined using a variety of available detection means, including nucleic acid hybridization and/or nucleic acid polymerization assays. For example, in some embodiments, the presence or absence of the mutations and/or SNPs are determined using a gene chip array, a TaqMan assay, or genomic DNA sequencing. The methodology may employ nucleic acid probes and/or primers designed for detecting an SNP or mutation described herein by any available assay format. Detection methods include, but are not limited to, Northern blot analysis, RNase protection, in situ methods, e.g. in situ hybridization, in vitro amplification methods (PCR, LCR, QRNA replicase or RNA-transcription/amplification (TAS, 3SR), reverse dot blot, and other detection assays that are known to those skilled in the art. Further, products obtained by in vitro amplification can be detected according to established methods, e.g. by separating the products on agarose or polyacrylamide gels and by subsequent staining with ethidium bromide or any other dye or reagent. Amplified products may be detected by using labeled primers or labeled dNTPs for amplification. The nucleic acid probes or primers may also be detectably labeled, for example, with a radioisotope, a bioluminescent compound, a chemiluminescent compound, a fluorescent compound, a metal chelate, or an enzyme.

[032] In addition to determining the presence or absence of mutational events in patient nucleic acids, the present invention further provides methods for evaluating IBD by determining the presence or absence of a protein variant, or the level of expression or activity of NELL1 in a biological sample from the patient. The biological sample is preferably from a tissue affected by the IBD, such as the colon, but may optionally be from any tissue expressing NELL1. When determining the presence or absence of a NELL1 variant in a biological sample, the invention involves contacting the sample, or material derived from the sample, with an antibody specific for a NELL1 variant that is associated with IBD, and observing or measuring an antibody-binding event. Antibodies specific for NELL1 variants are described more fully below. Alternatively, when determining the level of expression of

NELL1, any antibody recognizing NELL1 may be used, that is, such embodiments are not limited to the use of antibodies against NELL1 variants. Various antibody-based assays for measuring binding between the antibody and NELL1 variant are well known in the art.

NELL1 Polypeptides, Polynucleotide, and Antibodies

[033] In a second aspect, the invention provides novel variants of the NELL1 protein and encoding polynucleotides, as well as antibodies recognizing the novel NELL1 variants. Such products have use as diagnostic, prognostic, and therapeutic targets, as well as diagnostic, prognostic, and therapeutic agents for IBD, including CD.

[034] In accordance with this aspect, the present invention provides novel variants of the NELL1 protein and fragments thereof containing an amino acid mutation disclosed herein. The variants may include one or more of the mutations R136S, A153T, and/or R354W with respect to a mammalian (e.g., human, mouse, or rat) wild-type NELL-1 sequence, or a substantially homologous sequence containing the one or more mutations. Wild-type NELL1 sequences are defined by SEQ ID NOS: 1-3. In certain embodiments, the variant polypeptide further includes the mutation Q82R. These mutations are each associated with the presence of an IBD, and as such, the variant polypeptides find use as diagnostic and therapeutic targets, in conjunction with methods disclosed herein.

[035] This aspect further provides polynucleotides encoding the novel NELL1 variants of the invention, and complementary sequences. Such polynucleotides may be cloned into any suitable vector for replication of the polynucleotides, or expression of the variant polypeptides from promoter sequences, in host cells.

[036] The polynucleotides of the invention may be DNA, cDNA, synthetic DNA, synthetic RNA, or derivatives thereof. Such sequences may be isolated from genomic DNA (e.g., from a patient's cells), and therefore may optionally include naturally occurring introns, genic regions, nongenic regions, and regulatory regions. Alternatively, the polynucleotide may be isolated mRNA, or cDNA produced by reverse transcription, for example. In one embodiment, DNA containing all or part of the coding sequence for a NELL1 variant of the invention, is incorporated into a vector for expression of the encoded polypeptide in suitable host cells. For example, the NELL1 variant encoding sequence may be operably linked to a promoter sequence to drive expression of the NELL1 encoding RNA in a suitable host cell, including a bacterial host (e.g., *E. coli*), or eukaryotic host cell (e.g., yeast).

[037] The invention may employ any expression vector known in the art, including expression vectors derived from retroviruses, adenovirus, herpes or vaccinia viruses or from

various bacterial plasmids. Expression vectors may further be eukaryotic expression vectors, sufficient for delivery of the polynucleotide to organs, tissues or cell populations. Such techniques are well known in the art.

[038] The present invention further provides antibodies that recognize the variant polypeptide sequences of the invention, preferably in a specific fashion over the wild-type sequence. In some embodiments, the antibodies are produced against the variant polypeptide in an animal conventionally used for antibody production, or via an antibody library, which are well known in the art. Such antibodies may recognize a variant NELL1 protein of the invention 2-fold, five-fold, ten-fold, 100-fold, or more, better than the wild-type sequence. Such antibodies are useful as diagnostic and therapeutic agents for IBD, including Crohn Disease.

[039] Antibodies may be prepared by immunizing suitable mammalian hosts utilizing appropriate immunization protocols using the variant proteins of the invention or antigen-containing fragments thereof. To enhance immunogenicity, these proteins or fragments can be conjugated to suitable carriers. Methods for preparing immunogenic conjugates with carriers such as BSA, KLH or other carrier proteins are well known in the art. In some circumstances, direct conjugation using, for example, carbodiimide reagents may be effective; in other instances linking reagents such as those supplied by Pierce Chemical Co. (Rockford, IL) may be desirable to provide accessibility to the hapten. The hapten peptides can be extended at either the amino or carboxy terminus with a cysteine residue or interspersed with cysteine residues, for example, to facilitate linking to a carrier. Administration of the immunogens is conducted generally by injection over a suitable time period and with use of suitable adjuvants, as is generally understood in the art. During the immunization schedule, titers of antibodies are taken to determine adequacy of antibody formation. While the polyclonal antisera produced in this way may be satisfactory for some applications, for pharmaceutical compositions, use of monoclonal preparations is preferred. Immortalized cell lines which secrete the desired monoclonal antibodies may be prepared using standard methods, see e.g., Kohler & Milstein (1992) or modifications which affect immortalization of lymphocytes or spleen cells, as is generally known. The immortalized cell lines secreting the desired antibodies can be screened by immunoassay in which the antigen is the peptide hapten, polypeptide or protein. When the appropriate immortalized cell culture secreting the desired antibody is identified, the cells can be cultured either in vitro or by production in ascites fluid. The desired monoclonal antibodies may be recovered from the culture supernatant or from the ascites supernatant. Fragments of the monoclonal antibodies or the polyclonal antisera which contain the immunologically significant portion(s) can be used as antagonists, as well as the intact antibodies. Use of immunologically

reactive fragments, such as Fab or Fab' fragments, is often preferable, especially in a therapeutic context, as these fragments are generally less immunogenic than the whole immunoglobulin. The antibodies or fragments may also be produced, using current technology, by recombinant means. Antibody regions that bind specifically to the desired regions of the protein can also be produced in the context of chimeras derived from multiple species. Antibody regions that bind specifically to the desired regions of the protein can also be produced in the context of chimeras from multiple species, for instance, humanized antibodies. The antibody can therefore be a humanized antibody or a human antibody, as described in U.S. Patent 5,585,089 or Riechmann et al. (1988).

Diagnostic Kits

[040] The invention further provides a kit or array containing nucleic acid primers and/or probes for determining the presence and/or absence of IBD risk genotype in a patient sample. The kit may consist essentially of primers and/or probes related to evaluating an IBD genotype in a sample, and primers and/or probes related to necessary or meaningful assay controls. The kit for evaluating IBD may comprise nucleic acid probes and/or primers designed to detect ten or more SNPs associated with IBD, such as SNPs found in the genes listed in Table 1, and including the SNPs listed in Tables 1-5. Alternatively, the kit for evaluating IBD (e.g., for evaluating or determining an IBD genotype) may contain probes and/or primers for detecting at least one mutation in *NELL1* that is associated with IBD, or at least one mutation in the *5p13.1* locus that is associated with IBD, and optionally at least one mutation associated with an IBD in one or more of *NOD2*, the *5q31* locus, *DLG5*, *TNFSF15*, *ATG16L1*, *CARD4*, and *IL23R* (including SNPs described herein). In accordance with this aspect, the kit may be a companion diagnostic kit for evaluating IBD or determining an IBD genotype in a patient, and for selecting or predicting appropriate therapeutic intervention.

[041] In certain embodiments, the kit comprises probes and/or primers designed to detect the presence or absence of two or more SNPs selected from rs2076756 (#1 in Table 1), rs1992662 (#70 in Table 1), rs1992660 (#75 in Table 1), rs1793004 (#83 in Table 1), rs10521209 (#159 in Table 1), and rs2631372 (#163 in Table 1) in the patient sample. Such SNPs localize to the IBD-associated genes and loci of *NOD2, 5q31, 5p13.1*, and *NELL1.*

[042] The kits may comprise, or may further comprise, probes and/or primers designed to detect the presence or absence of a mutation in the *NELL1* gene in a biological sample, or a mutation in the corresponding 11p15.1 locus. The mutation(s) may include one or more SNPs listed in Table 1, 2, and 4 that localize to the *NELL1* gene, and which are associated with IBD, including the SNPs rs1793004 and rs951199, and/or may include rs8176785,

rs10500885, rs1158547, and rs1945404, and/or may include mutations encoding the Q82R, R136S, A153T, and/or R354W variants of NELL1.

[043] The kits may also comprise, or further comprise, probes and/or primers designed to detect the presence or absence of a mutation associated with IBD, in a patient's biological sample, in one or more of: the *5p13.1* locus (including SNPs listed in Tables 1, 3, and 5), including *PTGER4* (upstream); *ITGB6* (upstream); *GRM8* (downstream); *OR5V1* (downstream), *PPP3R2* (downstream); *NM_152575* (upstream); and *HNF4G* (intron). In these embodiments, the set of probes and/or primers may comprise probes and/or primers designed to detect one or more SNPs selected from rs1992662, rs1992660, rs1553575, rs7725523, rs2925757, rs6947579, rs10487428, rs10484545, rs4743484, rs7868736, and rs830772 in a biological sample.

[044] The kit may include a set of probes and/or primers designed to detect at least 10 IBD-associated polymorphisms (e.g., comprising or consisting essentially of IBD-associated genes disclosed herein), or may be designed to detect 20, 50, 100, 200, or more IBD-associated polymorphisms.

[045] In accordance with this aspect, the probes and primers may comprise antisense nucleic acids or oligonucleotides that are wholly or partially complementary to the diagnostic targets described herein. The probes and primers will be designed to detect the particular diagnostic target via an available nucleic acid detection assay format, which are well known in the art.

[046] In this context, the term "oligonucleotide" refers to naturally-occurring species or synthetic species formed from naturally-occurring subunits or their close homologs. The term may also refer to moieties that function similarly to oligonucleotides, but have non-naturally-occurring portions. Thus, oligonucleotides may have altered sugar moieties or inter-sugar linkages. Exemplary among these are phosphorothioate and other sulfur containing species which are known in the art. Such substitutions may comprise phosphorothioate bonds, methyl phosphonate bonds, or short chain alkyl or cycloalkyl structures. Oligonucleotides may also include species that include at least some modified base forms. Thus, purines and pyrimidines other than those normally found in nature may be so employed. Similarly, modifications on the furanosyl portions of the nucleotide subunits may also be effected. Examples of such modifications are 2'-O-alkyl- and 2'-halogen-substituted nucleotides. Some non-limiting examples of modifications at the 2' position of sugar moieties which may be useful include OH, SH, SCH_3, F, OCH_3, OCN, $O(CH_2)$, NH_2 and $O(CH_2)n$ CH_3, where n is from 1 to about 10. Such oligonucleotides are functionally interchangeable with natural oligonucleotides or synthesized oligonucleotides, which have

one or more differences from the natural structure. All such analogs are comprehended by this invention so long as they function effectively to hybridize with at least one diagnostic target of the invention.

[047] The oligonucleotides in accordance with this invention may comprise from about 3 to about 50 subunits or nucleotides. In some embodiments, oligonucleotides and analogs comprise from about 8 to about 25 subunits (e.g., nucleotides) and still more preferred to have from about 12 to about 20 subunits or nucleotides. As defined herein, a "subunit" is a base and sugar combination suitably bound to adjacent subunits through phosphodiester or other bonds.

[048] The kits of the invention may comprise probes and/or primers designed to detect the diagnostic targets via detection methods that include amplification, restriction enzyme cleavage, hybridization, sequencing, and cleavage.

[049] Amplification methods include: self sustained sequence replication (Guatelli et al., 1990), transcriptional amplification system (Kwoh et al., 1989), Q-Beta Replicase (Lizardi et al., 1988), isothermal amplification (e.g. Dean et al., 2002; and Hafner et al., 2001), or any other nucleic acid amplification method, followed by the detection of the amplified molecules using techniques well known to those of ordinary skill in the art. These detection schemes are especially useful for the detection of nucleic acid molecules if such molecules are present in very low number.

[050] Restriction enzyme cleavage methods include: isolating sample and control DNA, amplification (optional), digestion with one or more restriction endonucleases, determination of fragment length sizes by gel electrophoresis and comparing samples and controls. Differences in fragment length sizes between sample and control DNA indicates mutations in the sample DNA. Moreover, sequence specific ribozymes (see, e.g., U.S. Pat. No. 5,498,531 or DNAzyme e.g. U.S. Pat. No. 5,807,718) can be used to score for the presence of specific mutations by development or loss of a ribozyme or DNAzyme cleavage site.

[051] Hybridization methods include any measurement of the hybridization or gene expression levels, of sample nucleic acids to probes, and include microarray technology to detect several (e.g., more than 10, more than 100, or more than 1000) diagnostic targets. Thus, SNPs and mutations of the invention can be detected in a sample by hybridizing sample nucleic acids, e.g., DNA or RNA, to high density arrays or bead arrays containing oligonucleotide probes designed to hybridize thereto. Methods of forming high density arrays of oligonucleotides with a minimal number of synthetic steps are known. The oligonucleotide analogue array can be synthesized on a single or on multiple solid

substrates by a variety of methods, including, but not limited to, light-directed chemical coupling, and mechanically directed coupling.

[052] Nucleic acid hybridization simply involves contacting a probe and target nucleic acid under conditions where the probe and its complementary target can form stable hybrid duplexes through complementary base pairing. The nucleic acids that do not form hybrid duplexes are then washed away leaving the hybridized nucleic acids to be detected, typically through detection of an attached detectable label. It is generally recognized that nucleic acids are denatured by increasing the temperature or decreasing the salt concentration of the buffer containing the nucleic acids. Under low stringency conditions (e.g., low temperature and/or high salt) hybrid duplexes (e.g., DNA:DNA, RNA:RNA, or RNA:DNA) will form even where the annealed sequences are not perfectly complementary. Thus, specificity of hybridization is reduced at lower stringency. Conversely, at higher stringency (e.g., higher temperature or lower salt) successful hybridization tolerates fewer mismatches. One of skill in the art will appreciate that hybridization conditions may be selected to provide any degree of stringency.

[053] In a preferred embodiment, hybridization is performed at low stringency to ensure hybridization and then subsequent washes are performed at higher stringency to eliminate mismatched hybrid duplexes. Successive washes may be performed at increasingly higher stringency until a desired level of hybridization specificity is obtained. Stringency can also be increased by addition of agents such as formamide. Hybridization specificity may be evaluated by comparison of hybridization to the test probes with hybridization to the various controls that can be present (e.g., expression level control, normalization control, mismatch controls, etc.).

[054] In general, there is a tradeoff between hybridization specificity (stringency) and signal intensity. Thus, in a preferred embodiment, the wash is performed at the highest stringency that produces consistent results and that provides a signal intensity greater than approximately 10% of the background intensity. Thus, in a preferred embodiment, the hybridized array may be washed at successively higher stringency solutions and read between each wash. Analysis of the data sets thus produced will reveal a wash stringency above which the hybridization pattern is not appreciably altered and which provides adequate signal for the particular oligonucleotide probes of interest.

[055] Probes based on the sequences of the genes described above may be prepared by any commonly available method. Oligonucleotide probes for screening or assaying a tissue or cell sample are preferably of sufficient length to specifically hybridize only to appropriate, complementary genes or transcripts. Typically the oligonucleotide probes will be at least

about 10, 12, 14, 16, 18, 20 or 25 nucleotides in length. In some cases, longer probes of at least 30, 40, or 50 nucleotides will be desirable.

[056] The phrase “hybridizing specifically to” or “specifically hybridizes” refers to the binding, duplexing, or hybridizing of a molecule substantially to or only to a particular nucleotide sequence or sequences under stringent conditions when that sequence is present in a complex mixture (e.g., total cellular) DNA or RNA.

[057] As used herein a “probe” is defined as a nucleic acid, capable of binding to a target nucleic acid of complementary sequence through one or more types of chemical bonds, usually through complementary base pairing, usually through hydrogen bond formation. As used herein, a probe may include natural (i.e., A, G, U, C, or T) or modified bases (7-deazaguanosine, inosine, etc.). In addition, the bases in probes may be joined by a linkage other than a phosphodiester bond, so long as it does not interfere with hybridization. Thus, probes may be peptide nucleic acids in which the constituent bases are joined by peptide bonds rather than phosphodiester linkages.

[058] A variety of sequencing reactions known in the art can be used to directly sequence nucleic acids for the presence or the absence of one or more polymorphisms or mutations (such as those described herein). Examples of sequencing reactions include those based on techniques developed by Maxam and Gilbert (1977) or Sanger (1977). It is also contemplated that any of a variety of automated sequencing procedures can be utilized, including sequencing by mass spectrometry (see, e.g. PCT International Publication No. WO 94/16101; Cohen et al., 1996; and Griffin et al.,1993), real-time pyrophosphate sequencing method (Ronaghi et al.,1998; and Permutt et al., 2001) and sequencing by hybridization (see e.g. Drmanac et al., 2002).

[059] Other methods of detecting polymorphisms include methods in which protection from cleavage agents is used to detect mismatched bases in RNA/RNA, DNA/DNA or RNA/DNA heteroduplexes (Myers et al., 1985). In general, the technique of "mismatch cleavage" starts by providing heteroduplexes formed by hybridizing (labeled) RNA or DNA containing a wild-type sequence with potentially mutant RNA or DNA obtained from a sample. The double-stranded duplexes are treated with an agent who cleaves single-stranded regions of the duplex such as which will exist due to basepair mismatches between the control and sample strands. For instance, RNA/DNA duplexes can be treated with RNase and DNA/DNA hybrids treated with S1 nuclease to enzymatically digest the mismatched regions. In other embodiments, either DNA/DNA or RNA/DNA duplexes can be treated with hydroxylamine or osmium tetroxide and with piperidine in order to digest mismatched regions. After digestion of the mismatched regions, the resulting material is then separated by size on denaturing

polyacrylamide gels to determine the site of a mutation or SNP (see, for example, Cotton et al., 1988; and Saleeba et al., 1992). In a preferred embodiment, the control DNA or RNA can be labeled for detection.

[060] In still another embodiment, the mismatch cleavage reaction employs one or more proteins that recognize mismatched base pairs in double-stranded DNA (so called "DNA mismatch repair" enzymes) in defined systems for detecting and mapping polymorphisms. For example, the mutY enzyme of E. coli cleaves A at G/A mismatches (Hsu et al., 1994). Other examples include, but are not limited to, the MutHLS enzyme complex of E. coli (Smith and Modrich Proc. 1996) and Cel 1 from the celery (Kulinski et al., 2000) both cleave the DNA at various mismatches. According to an exemplary embodiment, a probe based on a polymorphic site is hybridized to a cDNA or other DNA product from a test cell or cells. The duplex is treated with a DNA mismatch repair enzyme, and the cleavage products, if any, can be detected from electrophoresis protocols or the like. See, for example, U.S. Pat. No. 5,459,039. Alternatively, the screen can be performed in vivo following the insertion of the heteroduplexes in an appropriate vector. The whole procedure is known to those ordinary skilled in the art and is referred to as mismatch repair detection (see e.g. Fakhrai-Rad et al., 2004).

[061] In other embodiments, alterations in electrophoretic mobility can be used to identify polymorphisms in a sample. For example, single strand conformation polymorphism (SSCP) analysis can be used to detect differences in electrophoretic mobility between mutant and wild type nucleic acids (Orita et al., 1989; Cotton et al., 1993; and Hayashi 1992). Single-stranded DNA fragments of case and control nucleic acids will be denatured and allowed to renature. The secondary structure of single-stranded nucleic acids varies according to sequence. The resulting alteration in electrophoretic mobility enables the detection of even a single base change. The DNA fragments may be labeled or detected with labeled probes. The sensitivity of the assay may be enhanced by using RNA (rather than DNA), in which the secondary structure is more sensitive to a change in sequence. In a preferred embodiment, the method utilizes heteroduplex analysis to separate double stranded heteroduplex molecules on the basis of changes in electrophoretic mobility (Kee et al., 1991).

[062] In yet another embodiment, the movement of mutant or wild-type fragments in a polyacrylamide gel containing a gradient of denaturant is assayed using denaturing gradient gel electrophoresis (DGGE) (Myers et al., 1985). When DGGE is used as the method of analysis, DNA will be modified to insure that it does not completely denature, for example by adding a GC clamp of approximately 40 bp of high-melting GC-rich DNA by PCR. In a further embodiment, a temperature gradient is used in place of a denaturing gradient to identify differences in the mobility of control and sample DNA (Rosenbaum et al., 1987). In

another embodiment, the mutant fragment is detected using denaturing HPLC (see e.g. Hoogendoorn et al., 2000).

[063] Examples of other techniques for detecting polymorphisms include, but are not limited to, selective oligonucleotide hybridization, selective amplification, selective primer extension, selective ligation, single-base extension, selective termination of extension or invasive cleavage assay. For example, oligonucleotide primers may be prepared in which the polymorphism is placed centrally and then hybridized to target DNA under conditions which permit hybridization only if a perfect match is found (Saiki et al., 1986; Saiki et al., 1989). Such oligonucleotides are hybridized to PCR amplified target DNA or a number of different mutations when the oligonucleotides are attached to the hybridizing membrane and hybridized with labeled target DNA. Alternatively, the amplification, the allele-specific hybridization and the detection can be done in a single assay following the principle of the 5' nuclease assay (e.g. see Livak et al., 1995). For example, the associated allele, a particular allele of a polymorphic locus, or the like is amplified by PCR in the presence of both allele-specific oligonucleotides, each specific for one or the other allele. Each probe has a different fluorescent dye at the 5' end and a quencher at the 3' end. During PCR, if one or the other or both allele-specific oligonucleotides are hybridized to the template, the Taq polymerase via its 5' exonuclease activity will release the corresponding dyes. The latter will thus reveal the genotype of the amplified product.

[064] Hybridization assays may also be carried out with a temperature gradient following the principle of dynamic allele-specific hybridization or like e.g. Jobs et al., (2003); and Bourgeois and Labuda, (2004). For example, the hybridization is done using one of the two allele-specific oligonucleotides labeled with a fluorescent dye, and an intercalating quencher under a gradually increasing temperature. At low temperature, the probe is hybridized to both the mismatched and full-matched template. The probe melts at a lower temperature when hybridized to the template with a mismatch. The release of the probe is captured by an emission of the fluorescent dye, away from the quencher. The probe melts at a higher temperature when hybridized to the template with no mismatch. The temperature-dependent fluorescence signals therefore indicate the absence or presence of an associated allele, a particular allele of a polymorphic locus, or the like (e.g. Jobs et al., 2003). Alternatively, the hybridization is done under a gradually decreasing temperature. In this case, both allele-specific oligonucleotides are hybridized to the template competitively. At high temperature none of the two probes are hybridized. Once the optimal temperature of the full-matched probe is reached, it hybridizes and leaves no target for the mismatched probe (e.g. Bourgeois and Labuda, 2004). In the latter case, if the allele-specific probes are differently labeled, then they are hybridized to a single PCR-amplified target. If the probes

are labeled with the same dye, then the probe cocktail is hybridized twice to identical templates with only one labeled probe, different in the two cocktails, in the presence of the unlabeled competitive probe.

[065] Alternatively, allele specific amplification technology that depends on selective PCR amplification may be used in conjunction with the present invention. Oligonucleotides used as primers for specific amplification may carry the associated allele, a particular allele of a polymorphic locus, or the like, also referred to as “mutation” of interest in the center of the molecule, so that amplification depends on differential hybridization (Gibbs et al., 1989) or at the extreme 3' end of one primer where, under appropriate conditions, mismatch can prevent, or reduce polymerase extension (Prossner, 1993). In addition it may be desirable to introduce a novel restriction site in the region of the mutation to create cleavage-based detection (Gasparini et al., 1992). It is anticipated that in certain embodiments, amplification may also be performed using Taq ligase for amplification (Barany, 1991). In such cases, ligation will occur only if there is a perfect match at the 3' end of the 5' sequence making it possible to detect the presence of a known associated allele, a particular allele of a polymorphic locus, or the like at a specific site by looking for the presence or absence of amplification. The products of such an oligonucleotide ligation assay can also be detected by means of gel electrophoresis. Furthermore, the oligonucleotides may contain universal tags used in PCR amplification and zip code tags that are different for each allele. The zip code tags are used to isolate a specific, labeled oligonucleotide that may contain a mobility modifier (e.g. Grossman et al., 1994).

[066] In yet another alternative, allele-specific elongation followed by ligation will form a template for PCR amplification. In such cases, elongation will occur only if there is a perfect match at the 3’ end of the allele-specific oligonucleotide using a DNA polymerase. This reaction is performed directly on the genomic DNA and the extension/ligation products are amplified by PCR. To this end, the oligonucleotides contain universal tags allowing amplification at a high multiplex level and a zip code for SNP identification. The PCR tags are designed in such a way that the two alleles of a SNP are amplified by different forward primers, each having a different dye. The zip code tags are the same for both alleles of a given SNPs and they are used for hybridization of the PCR-amplified products to oligonucleotides bound to a solid support, chip, bead array or like. For an example of the procedure, see Fan et al. (Cold Spring Harbor Symposia on Quantitative Biology, Vol. LXVIII, pp. 69-78 2003).

[067] Another alternative includes the single-base extension/ligation assay using a molecular inversion probe, consisting of a single, long oligonucleotide (see e.g. Hardenbol et al., 2003). In such an embodiment, the oligonucleotide hybridizes on both side of the SNP

locus directly on the genomic DNA, leaving a one-base gap at the SNP locus. The gap-filling, one-base extension/ligation is performed in four tubes, each having a different dNTP. Following this reaction, the oligonucleotide is circularized whereas unreactive, linear oligonucleotides are degraded using an exonuclease such as exonuclease I of E. coli. The circular oligonucleotides are then linearized and the products are amplified and labeled using universal tags on the oligonucleotides. The original oligonucleotide also contains a SNP-specific zip code allowing hybridization to oligonucleotides bound to a solid support, chip, and bead array or like. This reaction can be performed at a high multiplexed level.

[068] In another alternative, the associated allele, a particular allele of a polymorphic locus, or the like is scored by single-base extension (see e.g. U.S. Pat. No. 5,888,819). The template is first amplified by PCR. The extension oligonucleotide is then hybridized next to the SNP locus and the extension reaction is performed using a thermostable polymerase such as ThermoSequenase (GE Healthcare) in the presence of labeled ddNTPs. This reaction can therefore be cycled several times. The identity of the labeled ddNTP incorporated will reveal the genotype at the SNP locus. The labeled products can be detected by means of gel electrophoresis, fluorescence polarization (e.g. Chen et al., 1999) or by hybridization to oligonucleotides bound to a solid support, chip, and bead array or like. In the latter case, the extension oligonucleotide will contain a SNP-specific zip code tag.

[069] In yet another alternative, a SNP is scored by selective termination of extension. The template is first amplified by PCR and the extension oligonucleotide hybridizes in the vicinity of the SNP locus, close to but not necessarily adjacent to it. The extension reaction is carried out using a thermostable polymerase such as ThermoSequenase (GE Healthcare) in the presence of a mix of dNTPs and at least one ddNTP. The latter has to terminate the extension at one of the allele of the interrogated SNP, but not both such that the two alleles will generate extension products of different sizes. The extension product can then be detected by means of gel electrophoresis, in which case the extension products need to be labeled, or by mass spectrometry (see e.g. Storm et al., 2003).

[070] In another alternative, SNPs are detected using an invasive cleavage assay (see U.S. Pat. No. 6,090,543). There are five oligonucleotides per SNP to interrogate but these are used in a two step-reaction. During the primary reaction, three of the designed oligonucleotides are first hybridized directly to the genomic DNA. One of them is locus-specific and hybridizes up to the SNP locus (the pairing of the 3' base at the SNP locus is not necessary). There are two allele-specific oligonucleotides that hybridize in tandem to the locus-specific probe but also contain a 5' flap that is specific for each allele of the SNP. Depending upon hybridization of the allele-specific oligonucleotides at the base of the SNP locus, this creates a structure that is recognized by a cleavase enzyme (U.S. Pat. No.

6,090,606) and the allele-specific flap is released. During the secondary reaction, the flap fragments hybridize to a specific cassette to recreate the same structure as above except that the cleavage will release a small DNA fragment labeled with a fluorescent dye that can be detected using regular fluorescence detector. In the cassette, the emission of the dye is inhibited by a quencher.

Methods Of Treatment

[071] The present invention provides a method for treating IBD, including ulcerative colitis and Crohn Disease, by administering an effective amount of a NELL1 polypeptide or functional portion thereof, to a patient in need. Administration of the NELL1 polypeptide to patients suffering from IBD, or at risk of developing an IBD, may be effective to mitigate the effects of absent, partial inactivation, or abnormal expression of endogenous NELL1.

[072] The present invention provides methods of treating an IBD by expressing in vivo a NELL1 polynucleotide. These nucleic acids can be inserted into any of a number of well-known vectors for the transfection of target cells and organisms as described below. The nucleic acids are transfected into cells, ex vivo or in vivo, through the interaction of the vector and the target cell. The nucleic acids encoding a NELL1 protein or NELL1 sequence, under the control of a promoter, express the encoded protein, thereby mitigating the effects of absent, partial inactivation, or abnormal expression of endogenous NELL1.

[073] Alternatively, the invention provides nucleic acids, including expression constructs, that, when introduced into host cells expressing NELL1, express antisense and dsRNAs corresponding to portions of the NELL1 gene. Such expression of antisense and dsRNAs is effective to silence endogenous NELL1 expression via antisense or RNAi-mediated gene silencing.

[074] Such gene therapy procedures have been used to correct acquired and inherited genetic defects, cancer, and viral infection in a number of contexts. The ability to express artificial genes in humans facilitates the prevention and/or cure of many important human disorders, including many disorders which are not amenable to treatment by other therapies (for a review of gene therapy procedures, see Anderson, 1992; Nabel & Felgner, 1993; Mitani & Caskey, 1993; Mulligan, 1993; Dillon, 1993; Miller, 1992; Van Brunt, 1998; Vigne, 1995; Kremer & Perricaudet 1995; Doerfler & Bohm 1995; and Yu et al., 1994).

[075] Delivery of the gene or genetic material into the cell is the first critical step in gene therapy treatment of a disorder. A large number of delivery methods are well known to those of skill in the art. Preferably, the nucleic acids are administered for in vivo or ex vivo gene

therapy uses. Non-viral vector delivery systems include DNA plasmids, naked nucleic acid, and nucleic acid complexed with a delivery vehicle such as a liposome. Viral vector delivery systems include DNA and RNA viruses, which have either episomal or integrated genomes after delivery to the cell. For a review of gene therapy procedures, see the references included in the above section.

[076] The use of RNA or DNA based viral systems for the delivery of nucleic acids take advantage of highly evolved processes for targeting a virus to specific cells in the body and trafficking the viral payload to the nucleus. Viral vectors can be administered directly to patients (in vivo) or they can be used to treat cells in vitro and the modified cells are administered to patients (ex vivo). Conventional viral based systems for the delivery of nucleic acids could include retroviral, lentivirus, adenoviral, adeno-associated and herpes simplex virus vectors for gene transfer. Viral vectors are currently the most efficient and versatile method of gene transfer in target cells and tissues. Integration in the host genome is possible with the retrovirus, lentivirus, and adeno-associated virus gene transfer methods, often resulting in long term expression of the inserted transgene. Additionally, high transduction efficiencies have been observed in many different cell types and target tissues.

[077] The tropism of a retrovirus can be altered by incorporating foreign envelope proteins, expanding the potential target population of target cells. Lentiviral vectors are retroviral vectors that are able to transduce or infect non-dividing cells and typically produce high viral titers. Selection of a retroviral gene transfer system would therefore depend on the target tissue. Retroviral vectors are comprised of cis-acting long terminal repeats with packaging capacity for up to 6-10 kb of foreign sequence. The minimum cis-acting LTRs are sufficient for replication and packaging of the vectors, which are then used to integrate the therapeutic gene into the target cell to provide permanent transgene expression. Widely used retroviral vectors include those based upon murine leukemia virus (MuLV), gibbon ape leukemia virus (GaLV), Simian Immuno deficiency virus (SIV), human immuno deficiency virus (HIV), and combinations thereof (see, e.g., Buchscher et al., 1992; Johann et al., 1992; Sommerfelt et al., 1990; Wilson et al., 1989; Miller et al.,1999;and PCT/US94/05700).

[078] In applications where transient expression of the nucleic acid is preferred, adenoviral based systems are typically used. Adenoviral based vectors are capable of very high transduction efficiency in many cell types and do not require cell division. With such vectors, high titer and levels of expression have been obtained. This vector can be produced in large quantities in a relatively simple system. Adeno-associated virus ("AAV") vectors are also used to transduce cells with target nucleic acids, e.g., in the in vitro production of nucleic acids and peptides, and for in vivo and ex vivo gene therapy procedures (see, e.g., West et al., 1987; U.S. Pat. No. 4,797,368; WO 93/24641; Kotin, 1994; Muzyczka, 1994).

Construction of recombinant AAV vectors is described in a number of publications, including U.S. Pat. No. 5,173,414; Tratschin et al., 1985; Tratschin, et al., 1984; Hermonat & Muzyczka, 1984; and Samulski et al., 1989.

[079] In particular, numerous viral vector approaches are currently available for gene transfer in clinical trials, with retroviral vectors by far the most frequently used system. All of these viral vectors utilize approaches that involve complementation of defective vectors by genes inserted into helper cell lines to generate the transducing agent. pLASN and MFG-S are examples are retroviral vectors that have been used in clinical trials (Dunbar et al., 1995; Kohn et al., 1995; Malech et al., 1997). PA317/pLASN was the first therapeutic vector used in a gene therapy trial (Blaese et al., 1995). Transduction efficiencies of 50% or greater have been observed for MFG-S packaged vectors (Ellem et al., 1997; and Dranoff et al., 1997).

[080] Recombinant adeno-associated virus vectors (rAAV) are a promising alternative gene delivery systems based on the defective and nonpathogenic parvovirus adeno-associated type 2 virus. All vectors are derived from a plasmid that retains only the AAV 145 bp inverted terminal repeats flanking the transgene expression cassette. Efficient gene transfer and stable transgene delivery due to integration into the genomes of the transduced cell are key features for this vector system (Wagner et al., 1998, Kearns et al., 1996).

[081] Replication-deficient recombinant adenoviral vectors (Ad) are predominantly used in transient expression gene therapy; because they can be produced at high titer and they readily infect a number of different cell types. Most adenovirus vectors are engineered such that a transgene replaces the Ad E1a, E1b, and E3 genes; subsequently the replication defector vector is propagated in human 293 cells that supply the deleted gene function in trans. Ad vectors can transduce multiple types of tissues in vivo, including nondividing, differentiated cells such as those found in the liver, kidney and muscle tissues. Conventional Ad vectors have a large carrying capacity. An example of the use of an Ad vector in a clinical trial involved polynucleotide therapy for antitumor immunization with intramuscular injection (Sterman et al., 1998). Additional examples of the use of adenovirus vectors for gene transfer in clinical trials include Rosenecker et al., 1996; Sterman et al., 1998; Welsh et al., 1995; Alvarez et al., 1997; Topf et al., 1998.

[082] Packaging cells are used to form virus particles that are capable of infecting a host cell. Such cells include 293 cells, which package adenovirus, and PA317 cells, which package retrovirus. Viral vectors used in gene therapy are usually generated by a producer cell line that packages a nucleic acid vector into a viral particle. The vectors typically contain the minimal viral sequences required for packaging and subsequent integration into a host,

other viral sequences being replaced by an expression cassette for the protein to be expressed. The missing viral functions are supplied in trans by the packaging cell line. For example, AAV vectors used in gene therapy typically only possess ITR sequences from the AAV genome which are required for packaging and integration into the host genome. Viral DNA is packaged in a cell line, which contains a helper plasmid encoding the other AAV genes, namely rep and cap, but lacking ITR sequences. The cell line is also infected with adenovirus as a helper. The helper virus promotes replication of the AAV vector and expression of AAV genes from the helper plasmid. The helper plasmid is not packaged in significant amounts due to a lack of ITR sequences. Contamination with adenovirus can be reduced by, e.g., heat treatment to which adenovirus is more sensitive than AAV.

[083] In many gene therapy applications, it is desirable that the gene therapy vector be delivered with a high degree of specificity to a particular tissue type. A viral vector is typically modified to have specificity for a given cell type by expressing a ligand as a fusion protein with a viral coat protein on the viruses outer surface. The ligand is chosen to have affinity for a receptor known to be present on the cell type of interest. For example, Han et al., 1995, reported that Moloney murine leukemia virus can be modified to express human heregulin fused to gp70, and the recombinant virus infects certain human breast cancer cells expressing human epidermal growth factor receptor. This principle can be extended to other pairs of viruses expressing a ligand fusion protein and target cells expressing a receptor. For example, filamentous phage can be engineered to display antibody fragments (e.g., Fab or Fv) having specific binding affinity for virtually any chosen cellular receptor. Although the above description applies primarily to viral vectors, the same principles can be applied to nonviral vectors. Such vectors can be engineered to contain specific uptake sequences thought to favor uptake by specific target cells.

[084] Gene therapy vectors can be delivered in vivo by administration to an individual patient, typically by systemic administration (e.g., intravenous, intraperitoneal, intramuscular, subdermal, or intracranial infusion) or topical application. Alternatively, vectors can be delivered to cells ex vivo, such as cells explanted from an individual patient (e.g., lymphocytes, bone marrow aspirates, and tissue biopsy) or universal donor hematopoietic stem cells, followed by reimplantation of the cells into a patient, usually after selection for cells which have incorporated the vector.

[085] Ex vivo cell transfection for diagnostics, research, or for gene therapy (e.g., via re-infusion of the transfected cells into the host organism) is well known to those of skill in the art. In a preferred embodiment, cells are isolated from the subject organism, transfected with a nucleic acid (gene or cDNA), and re-infused back into the subject organism (e.g., patient). Various cell types suitable for ex vivo transfection are well known to those of skill in the art

(see, e.g., Freshney et al., 1994; and the references cited therein for a discussion of how to isolate and culture cells from patients).

[086] In one embodiment, stem cells are used in ex vivo procedures for cell transfection and gene therapy. The advantage to using stem cells is that they can be differentiated into other cell types in vitro, or can be introduced into a mammal (such as the donor of the cells) where they will engraft in the bone marrow.

[087] Stem cells are isolated for transduction and differentiation using known methods. For example, stem cells are isolated from bone marrow cells by panning the bone marrow cells with antibodies which bind unwanted cells, such as CD4+ and CD8+ (T cells), CD45+ (panB cells), GR-1 (granulocytes), and lad (differentiated antigen presenting cells).

[088] Vectors (e.g., retroviruses, adenoviruses, liposomes, etc.) containing therapeutic nucleic acids can be also administered directly to the organism for transduction of cells in vivo. Alternatively, naked DNA can be administered.

[089] Administration is by any of the routes normally used for introducing a molecule into ultimate contact with blood or tissue cells, as described above. The nucleic acids are administered in any suitable manner, preferably with the pharmaceutically acceptable carriers described above. Suitable methods of administering such nucleic acids are available and well known to those of skill in the art, and, although more than one route can be used to administer a particular composition, a particular route can often provide a more immediate and more effective reaction than another route (see Samulski et al., 1989). The present invention is not limited to any method of administering such nucleic acids, but preferentially uses the methods described herein.

[090] In other embodiments, the invention provides methods of treating IBD, by administering antibodies, including synthetic antibodies and antibody fragments, specific for NELL1 to patients in need of treatment. The antibodies are generally administered in amounts effective to inhibit NELL1 receptor or ligand binding. For example, such antibodies may block NELL1 receptor binding in vivo, thereby restoring normal levels NELL1 activity. The antibodies may be directed, for example, to the EGF or TSPN domains of NELL1. Alternatively, mimetics of NELL1 may be prepared, including antidiotypic antibodies, effective to act as agonists at the NELL1 receptor. Various forms of antibodies sufficient for these purposes are described elsewhere herein.

Screening Assays

[091] The invention further provides methods of screening for agonists and antagonists, and compounds or agents that modulate the expression, of NELL1. Such compounds and agents find use in the treatment of IBD, including the development and manufacture of treatments for IBD.

[092] In this aspect, the invention comprises contacting the NELL1 polypeptide, for example, as expressed in a suitable host cell, or as present in a suitable in vitro system, with a test compound or agent. The level of NELL1 activity (as described in the art, and as measured via any suitable NELL1 assay described in the art), or in some cases the level of NELL1 expression, may then be compared to controls to identify an agonist or antagonist of NELL1 (or an up- or down-regulator as the case may be). Controls may include positive controls for NELL1 activation or expression, and appropriate negative controls.

[093] In certain embodiments, the method may employ the NELL1 variants described herein, in order to identify a suitable agonist or antagonist of the NELL1 variant.

[094] Agents that are assayed in the above method can be randomly selected or rationally selected or designed. As used herein, an agent is said to be randomly selected when the agent is chosen randomly without considering the specific sequences involved in the association of the protein of the invention alone or with its associated substrates, binding partners, etc. An example of randomly selected agents is the use of a chemical library or a peptide combinatorial library, or a growth broth of an organism. As used herein, an agent is said to be rationally selected or designed when the agent is chosen on a non-random basis which takes into account the sequence of the target site or its conformation in connection with the agent's action. Agents can be rationally selected or rationally designed by utilizing the peptide sequences that make up these sites. For example, a rationally selected peptide agent can be a peptide whose amino acid sequence is identical to or a derivative of any functional consensus site. The agents of the present invention can be, as examples, oligonucleotides, antisense polynucleotides, interfering RNA, peptides, peptide mimetics, antibodies, antibody fragments, small molecules, vitamin derivatives, as well as carbohydrates. Peptide agents of the invention can be prepared using standard solid phase (or solution phase) peptide synthesis methods, as is known in the art. In addition, the DNA encoding these peptides may be synthesized using commercially available oligonucleotide synthesis instrumentation and produced recombinantly using standard recombinant production systems. The production using solid phase peptide synthesis is necessitated if non-gene-encoded amino acids are to be included.

[095] Another class of agents useful in this aspect includes antibodies or fragments thereof that bind to NELL1 or a variant thereof. Antibody agents can be obtained by immunization of suitable mammalian subjects with peptides, containing as antigenic regions, those portions of the protein intended to be targeted by the antibodies (see section above of antibodies as probes for standard antibody preparation methodologies).

[096] In yet another class of agents, the present invention includes peptide mimetics that mimic the three-dimensional structure of NELL1 or a variant thereof. Such peptide mimetics may have significant advantages over naturally occurring peptides, including, for example: more economical production, greater chemical stability, enhanced pharmacological properties (half-life, absorption, potency, efficacy, etc.), altered specificity (e.g., a broad-spectrum of biological activities), reduced antigenicity and others. In one form, mimetics are peptide-containing molecules that mimic elements of protein secondary structure. The underlying rationale behind the use of peptide mimetics is that the peptide backbone of proteins exists chiefly to orient amino acid side chains in such a way as to facilitate molecular interactions, such as those of antibody and antigen. A peptide mimetic is expected to permit molecular interactions similar to the natural molecule. In another form, peptide analogs are commonly used in the pharmaceutical industry as non-peptide drugs with properties analogous to those of the template peptide. These types of non-peptide compounds are also referred to as peptide mimetics or peptidomimetics (Fauchere, 1986; Veber & Freidinger, 1985; Evans et al., 1987) which are usually developed with the aid of computerized molecular modeling. Peptide mimetics that are structurally similar to therapeutically useful peptides may be used to produce an equivalent therapeutic or prophylactic effect. Generally, peptide mimetics are structurally similar to a paradigm polypeptide (i.e., a polypeptide that has a biochemical property or pharmacological activity), but have one or more peptide linkages optionally replaced by a linkage using methods known in the art. Labeling of peptide mimetics usually involves covalent attachment of one or more labels, directly or through a spacer (e.g., an amide group), to non-interfering position(s) on the peptide mimetic that are predicted by quantitative structure-activity data and molecular modeling. Such non-interfering positions generally are positions that do not form direct contacts with the macromolecule(s) to which the peptide mimetic binds to produce the therapeutic effect. Derivitization (e.g., labeling) of peptide mimetics should not substantially interfere with the desired biological or pharmacological activity of the peptide mimetic. The use of peptide mimetics can be enhanced through the use of combinatorial chemistry to create drug libraries. The design of peptide mimetics can be aided by identifying amino acid mutations that increase or decrease binding of the protein to its binding partners. Approaches that can be used include the yeast two hybrid method (see

Chien et al., 1991) and the phage display method. The two hybrid method detects protein-protein interactions in yeast (Fields et al., 1989). The phage display method detects the interaction between an immobilized protein and a protein that is expressed on the surface of phages such as lambda and M13 (Amberg et al., 1993; Hogrefe et al., 1993). These methods allow positive and negative selection for protein-protein interactions and the identification of the sequences that determine these interactions.

EXAMPLES

METHODS

Patient Recruitment

[097] German patients and controls in panels A, B, and C partially overlap with samples that have been used in other studies before [12,15,65,66]. Panels A and B almost completely overlap with the panels (also termed panel A and B) that were used in a recently published IBD association screen of non-synonymous SNPs [15]. In this non-synonymous SNP scan no coding SNPs that were evaluable were located in *NELL1*. All patients were recruited at the Charite´ University Hospital (Berlin, Germany) and the Department of General Internal Medicine of the Christian-Albrechts-University (Kiel, Germany), with the support of the German Crohn and Colitis Foundation. Clinical, radiological and endoscopic (i.e. type and distribution of lesions) examinations were required to unequivocally confirm the diagnosis of Crohn disease or ulcerative colitis [67,68], and histological findings also had to be confirmative of, or compatible with, the diagnosis. In the case of uncertainty, patients were excluded from the study. German control individuals were obtained from the POPGEN biobank [25].

[098] The UK patients (Panel E) were recruited as described [69]; UK controls were obtained from the 1958 British Birth Cohort.

[099] The French/Canadian trios (Panel D) were sampled from the Quebec founder population (QFP). Membership of the founder population was defined as having four grandparents with French Canadian family names who were born in the Province of Quebec, Canada, or in adjacent areas of the Provinces of New Brunswick and Ontario, or in New England or New York State (USA).

[0100] Informed written consent was obtained from all study participants. All collection protocols were approved by the institutional review committees of the participating centres.

SNP Genotyping with the Affymetrix 100k Gene Chip Array

[0101] Genotyping of cases and controls was carried out using the Affymetrix GeneChip® Human Mapping 50K Xba and Hind Arrays (Affymetrix, Santa Clara, CA, USA). Genotypes were called by the GeneChip® DNA Analysis Software (GDAS v2.0, Affymetrix). Gender was verified by counting the heterozygous SNPs on the X chromosome. Quality checks further comprised the verification of individual sample call rates (≥90%) and, to ensure that no samples were confused, the 31 identical SNPs present on both chips were checked for identical genotypes for the same individual. SNPs that had a low genotyping success rate (<90%), were monomorphic, or deviated from Hardy-Weinberg equilibrium ($p \leq 0.01$) were eliminated from subsequent analyses. Experimental details concerning the genotyping of the 100k SNP set are provided in Matsuzaki et al. [22].

Follow-up Genotyping and Sequencing

[0102] SNPlex™ (Applied Biosystems, Foster City, CA, USA) genotyping of panels A, B, and C was carried out as recently described [15]. Genotype concordance rates for SNPs rs1793004, rs1992660, and rs1992662 were checked using TaqMan (Applied Biosystems) as an independent genotyping technology on an automated platform [70] and the functionally tested assays C____392093_10, C__11472026_10, and C__11472042_10. All three concordance rates were >98%, excluding genotyping errors as a potential source of false-positive associations. The same three TaqMan assays were also used to genotype panels C and E. Genotypes for panel D were generated at Genizon BioSciences using the Illumina GoldenGate™ platform (Illumina, San Diego, CA, USA). All process data were logged into, and administered by, a database-driven LIMS [71]. TaqMan genotyping of *NOD2*/Arg702Trp, *NOD2*/Gly908Arg, *NOD2*/Leu1007fs, *DLG5*/Arg30Gln, *DLG5*/Pro1371Gln, *DLG5*/e26, and *ATG16L1*/Thr300Ala was performed using previously described assays [12,15,72]. *IL23R*/ Arg381Gln, *NELL1*/rs8176785, and *NELL1*/rs8176786 were genotyped in panels A, B, C, and E using the functionally tested assay C___1272298_10, C___3203197_10, and C__ 32647553_10, respectively (Applied Biosystems). Prior to statistical analyses, the same cut-off criteria as described above for the 100k analysis ($p_{HWE}>0.01$, $MAF_{controls}>0$, callrate≥90%) were applied to the SNPs under study.

[0103] Sequencing of genomic DNA was performed using Applied Biosystems BigDye™ chemistry according to the supplier's recommendations. Traces were inspected for the presence of SNPs and InDels using novoSNP [73].

Statistical Analysis

[0104] Genome-wide data analysis was carried out using an updated version of GENOMIZER [74]. Association hits that passed the quality criteria were extracted using the ''GenomizerHits'' tool. Haploview 4.0 [75] was used for association analysis, transmission disequilibrium tests, and LD quantification of the replication data. Fisher's exact test was used when appropriate. The supplementary p-value plots and quantile-quantile plots were created using R. Single-marker disease associations and possible marker-marker interactions were assessed for statistical significance by means of logistic regression analysis (forward selection), as implemented in the procedure LOGISTIC of the SAS software package (SAS Institute, Cary NC, USA). Haplotype analyses were carried out using COCAPHASE 2.403 [76] and PHASE 2.1 [77,78].

RT-PCR, Western Blot and Immunohistochemistry

[0105] For the assessment of tissue-specific expression patterns, a commercial tissue panel was employed (Clontech, Palo Alto, CA, USA). Primers used for amplification of *NELL1* were (NELL1_14- 16_F ACCTTCCTGGGTTATATCGCTGTG (SEQ ID NO: 10) and NELL1_14- 16_R TCTCGCAGTGGCTTCCTGTG (SEQ ID NO: 11), expected amplicon length: 285 bp). The following conditions were applied: denaturation for 5 min at 95°C; 40 cycles of 30 sec at 95°C, 20 sec at 60°C, 45 sec at 72°C; final extension for 10 min at 72°C. To confirm the use of equal amounts of RNA in each experiment, all samples were checked in parallel for β-actin mRNA expression. All amplified DNA fragments were analyzed on 2% agarose gels and subsequently documented by a BioDoc Analyzer (Biometra, Göttingen, Germany).

[0106] Paraformaldehyde-fixed paraffin-embedded biopsies from normal controls (n= 6) and from patients with confirmed colonic CD (n = 6) were analysed. Two slides of each biopsy were stained with hematoxylin-eosin for routine histological evaluation. The other slides were subjected to a citrate-based antigen retrieval procedure, permeabilized by incubation with 0.1% Triton X-100 in 0.1M phosphate-buffered saline (PBS), washed three times in PBS and blocked with 0.75% bovine serum albumin in PBS for 20 minutes. Sections were subsequently incubated with the primary antibody (anti-NELL1, Abnova , mouse monoclonal) at a 1:500 dilution in 0.75% BSA overnight at 4°C. After washing in PBS, tissue-bound antibody was detected using biotinylated goat-anti mouse (Vector Laboratory, Burlingame, CA) followed by HRP-conjugated avidin, both diluted at 1:100 in PBS. Controls were included using irrelevant primary antibodies as well as omitting the primary antibodies using only secondary antibodies and/or HRP-conjugated avidin. No significant staining was observed with any of these controls (data not shown). Bound antibody was detected by

standard chromogen technique (Vector Laboratory) and visualized by an Axiophot microscope (Zeiss, Jena, Germany). Pictures were captured by a digital camera system (Axiocam, Zeiss).

[0107] Western blot analysis was performed as described [79]. In brief, 20 µg of protein lysates freshly derived from colonic biopsies of four healthy controls without any obvious intestinal pathology and four CD patients with confirmed ileal and colonic inflammation were lysed, separated by SDS polyacrylamide gel electrophoresis and transferred to PVDF membrane by standard techniques. NELL1 was detected using the same monoclonal anti-NELL1 antibody also employed for immunohistochemistry.

In silico Protein Analysis

[0108] Aligned sequences were retrieved from the UniProt database and protein domain architectures taken from the NCBI conserved domain search website. To predict the 3D structure of the N-terminal domain of NELL1, we explored the fold recognition results returned by the web servers GenTHREADER and FFAS03. Based upon the very similar server predictions, a pair-wise sequence-structure alignment of NELL1 to the crystal structure of the human thrombospondin-1 N-terminal domain (TSPN) was constructed as input for the 3D-modeling server WHATIF, which returned a structure model of the NELL1 N-terminal domain (**Figure 4**).

Genome-wide association scan

[0109] A total of 116,161 SNPs were genotyped in case-control panel A. Of these, 92,387 SNPs had a call rate ≥ 90%, were polymorphic in panel A, and showed no significant departure from Hardy-Weinberg equilibrium ($p_{HWE} \leq 0.01$ in controls). At an unadjusted per-test significance level of 5%, the experiment had 80% power to detect an odds ratio of 1.6, and 33% power to detect an odds ratio of 1.3, assuming that 20% of the controls were carriers of the risk factor. The GWS results were not corrected for potential population substructure because (i) very low ($<10^{-3}$) F_{ST} values have previously been reported for different geographic regions of Germany [24], (ii) patients of panel A were all selected from the Northern part of Germany, and were therefore geographically matched to the population-representative controls from the POPGEN biobank [25], (iii) quantile-quantile plots, which can help to identify spurious association results [26], revealed no inflation of the X^2 statistics, and (iv) replication criteria included confirmation by family-based association tests (transmission disequilibrium test, TDT), which are robust against population stratification [27].

Replication

[0110] The 200 most significant SNPs in the GWS were next genotyped in two independent German samples. "Replication" was considered to have been achieved if the p-values of both, the case-control analysis and the family-based TDT were < 5%. Replication in two independent samples also rendered test-wise Bonferroni correction superfluous, which would have been overly conservative in a replication setting anyway [28]. In addition to rs2631372 (#163), which localizes to the *5q31* haplotype [10], an association with CD was confirmed for rs2076756 (#1, $p_{CCA} < 10^{-12}$) and rs10521209 (#159) in *NOD2* [7–9]. The recently reported *5p13.1* locus [18] was also replicated (rs1992660/#70, rs1992662/#75), and a novel susceptibility gene, *NELL1*, was identified (rs1793004/#83). While only these six SNPs were found to be nominally significant in both, the TDT and the case-control analysis, and therefore fulfilled the formal replication criteria, 47 SNPs were significant in only one test, including two more SNPs in the *5p13.1* region (#79, #105), one in *NELL1* (#116), and one in the *IBD5* region (#191). In view of the low power of the TDT, it appears worth mentioning that use of p_{CCA} or $p_{CCG} \leq 10^{-2}$ as the sole replication criterion would have led to the additional acceptance of rs2925757 (*ITGB6*, upstream), rs6947579 (*GRM8*, downstream), rs10484545 (*OR5V1*, downstream), rs4743484 (*PPP3R2*, downstream), rs7868736 (NM_152575, upstream), and rs830772 (*HNF4G*, intron) as confirmed associations.

[0111] We did not detect our previously reported CD associations of *ATG16L1* [15], and *DLG5* [12] in this screening and did not see any association with and *IL23R* [16]. However, SNP coverage around these genes was very low. In order to benchmark our experiments, relevant SNPs in these genes were therefore genotyped in panels A and B, using TaqMan technology, and a disease association was observed for SNPs in all three genes. Interestingly, haplotype A-tagging SNP e26 in the *DLG5* gene was replicated (over-transmission of common allele T, T:U=275:219, p=0.012), while the associations of non-synonymous SNPs Arg30Gln and Pro1371Gln did not reach statistical significance.

[0112] To corroborate our main association findings, we examined the significantly associated *NELL1* and *5p13.1* SNPs in two additional, independent Caucasian CD samples: Panel D, which comprised population-based Falk-Rubinstein trios from the Quebec founder population (QFP), and panel E, a case-control sample from the UK. The *NELL1* association was replicated in the QFP (over-transmission of the common C allele, T:U=140:107, p=0.036) sample. In addition, the association of *5p13.1* SNP 1992660 was replicated in the QFP case-control sample (p=0.0081) and the combined p-value for panels B, D, and E was 1.24×10^{-7} in an allele-based test. The odds ratio for homozygosity of the common A allele was 1.36 (95% CI: 1.36–2.04). In the UK sample (panel E), the associations of SNPs

rs1992660 and rs1992662 were replicated with p-values (allelic X^2 test) of 0.036 and 0.0011, respectively, while the *NELL1* SNP association did not achieve formal significance.

Evaluation in Ulcerative Colitis (UC)

[0113] The three SNPs rs1793004 (*NELL1*), rs1992660, and rs1992662 (both *5p13.1*) with a confirmed CD association were also analysed in a German UC panel (panel C, 1059 single patients and 419 trios). The *NELL1* SNP rs1793004 also showed a disease association in the UC case-control panel (p=0.0017 in the allele-based X^2 test) and the odds ratio for homozygosity of the common C allele was 1.54 (95% CI: 1.08–2.20). Given the similar odds ratio in UC and CD (1.76; 95% CI: 1.27–2.45), *NELL1* appears to be a ubiquitous IBD susceptibility gene (combined $p<10^{-6}$; OR=1.66, 95% CI:1.30–2.11). No association to UC was detected for the *5p13.1* locus.

Fine mapping around NELL1

[0114] Fine mapping around the *NELL1* gene was carried out in replication panels B and D using HapMap tagging SNPs at a density of 8 kb. Twenty-one SNPs in the *NELL1* gene yielded a p-value < 0.05 in the single-point analyses of panel B (12 in panel D), of which several markers were significant in both, the TDT and case-control test. Results were not corrected for multiple testing because the association between CD and the *NELL1* locus was regarded as established through the previous analyses of panels A and B.

[0115] *NELL1* comprises several regions of increased recombination, scattered over a total of 906 kb. Disease associations were found with various small linkage disequilibrium (LD) blocks, suggesting the existence of more than one causal variant in the gene. In a logistic regression analysis of the combined panels A+B, the best model fit was achieved with SNPs rs1793004, rs951199, rs8176785, rs10500885, rs1158547, and rs1945404. The main association peak was located 5' of the gene, although a few significant associations were also found towards the 3' end. The signal sharply declines 5' of the *NELL1* gene, thereby excluding an involvement of the proximate *SLC6A5* gene. A gender-stratified analysis (data not shown) of the 117 SNPs in panel B confirmed a disease association in both genders.

Detection of Additional DNA Sequence Variants

[0116] Since rs1793004 clearly localizes to *NELL1*, a systematic search for additional, potentially disease-associated variants in the gene was carried out by re-sequencing all exons, splice sites, and the promoter region in 47 CD patients. Apart from verifying 26 already annotated variants, five new polymorphisms were identified, two of which represented novel non-synonymous SNPs (nsSNPs): NELL01_02 (R136S) and NELL01_03 (A153T). Both nsSNPs were located in exon 4 and mapped to the thrombospondin N-

terminal domain (TSPN) of the NELL1 protein. Two known, common nsSNPs were verified among the 26 annotated SNPs, namely rs8176785 (Q82R) in exon 3 and rs8176786 (R354W) in exon 10. Variant Q82R is located in the TSPN domain, while R354W resides in a von Willebrand factor type C (VWC) domain. *In-silico* analysis, including multiple sequence alignment of NELL homologues and structural modeling of the TSPN domain, revealed a strong conservation of the variant positions **(Figures 2 and 3)**.

[0117] The novel nsSNPs were too rare to warrant formal statistical analysis (total occurrence of heterozygotes in panel B (CD/ controls): 2/0 for R136S and 10/9 for A153T). While common nsSNP rs8176786 showed a disease association in panel E (p = 0.048), the second common nsSNP, rs8176785, was significantly associated with CD in panel B (p= 0.039), and with UC in panel C (p = 0.013). The combined p-value in the full German IBD sample (A+B+C) was 0.0048 in a genotype-based X^2 test (2 degrees of freedom).

Expression and Localization of NELL1 within the Intestinal Mucosa

[0118] When *NELL1* transcript levels were investigated by RT-PCR in a tissue panel, high expression became apparent in small intestine, kidney, prostate, and brain, whereas moderate expression was seen in colonic mucosa and in immune-relevant organs/cells such as thymus and spleen (**Figure 1A**). The localization of *NELL1* in the colonic mucosa was investigated by immunohistochemistry (**Figures 1B and 1C**). Immunoreactivity was confined to large mononuclear cells in the lamina propria. In Western blot experiments using colonic biopsy specimen from normal controls and CD patients (**Figure 1E**), the antibody recognized a single 90 kDa band corresponding to the correct size of the annotated NELL1 transcript (AK127805). Real-time quantitative PCR revealed no significant difference between normal and patient tissue.

Fine Mapping of 5p13.1

[0119] The 650 kb susceptibility region on 5p13.1, upstream of the PTGER4 gene, was subjected to fine mapping in panels B (1 SNP/ 24 kb) and D (1 SNP/3 kb). Several SNPs showed a consistent disease association in both panels. The strongest effect in the combined case-control panel (A+B) was seen for SNP rs1553575 (odds ratio for homozygotes of the common G allele: 1.78; 95% CI: 1.32–2.40).

[0120] Interestingly, the gender-stratified analysis of 5p13.1 SNPs showed that the association signal was stronger in females than in males, suggesting that females carrying the predisposing allele(s) of this locus are at higher risk to develop Crohn disease than their male counterparts. To have comparable power, the same number of male and female individuals were randomly drawn from the combined panel (378 controls, 343 cases).

Interaction with Known Disease Loci

[0121] Logistic regression analysis and a Breslow-Day test for odds ratio heterogeneity were used to analyse the full German case-control panel (A+B) for potential epistatic effects. No statistically significant interactions were observed, neither between polymorphisms within the *NELL1* gene (rs1793004) or the *5p13.1* region (rs1992660 and rs1992662), nor between these loci and any of the known disease-associated variants in *IL23R* (rs11209026/ Arg381Gln), *NOD2* (rs2066844/Arg702Trp, rs2066845/ Gly908Arg, rs2066847/Leu1007fs), *ATG16L1* (rs2241880/ T300A), *DLG5* (rs1248696/Arg30Gln), or in the *IBD5* region (tagging SNP IGR2063_b1 [11,29]).

DISCUSSION

[0122] We have identified *NELL1* as a novel disease gene for Crohn disease (CD), a result that was obtained in a genome-wide casecontrol association scan with 116,161 SNPs and by extensive replication in three independent samples from three distinct ethnicities. In a recently published GWAS from the UK population [30,31] (1,748 CD patients and 2,938 controls genotyped), the *NELL1* region was covered with 263 SNPs. Of these, 23 SNPs were significantly associated with CD ($p<0.05$ under an additive or general model) and six SNPs had a $p<0.01$: rs7122630, rs4475916, rs7115151, rs11025862, rs2063913, rs11026037. The *NELL1* region was not subjected to replication in the UK scan, since none of the 23 SNPs fell below the chosen cut-off ($p<010^{-5}$).

[0123] In addition to identifying *NELL1* as a CD risk factor, we also replicated the disease association recently described for the *5p13.1* region [18]. The genome-wide scan also confirmed two of the previously known CD loci, *NOD2* and *5q31*, but it should be pointed out that the marker set only covered 31% of the genome [32,33]. The previously established disease association of neither *IL23R* [17], nor *DLG5* [12], nor *ATG16L1* [15], were detected. However, not a single SNP, for example, in the *ATG16L1* gene was present on the Affymetrix GeneChip® Human Mapping 100K Set and coverage of all these genes was low. Targeted *post-hoc* genotyping of the relevant SNPs in the German screening and replication panel was therefore carried out and confirmed the CD associations of *ATG16L1*, *IL23R* and *DLG5* in our study sample. We replicated the association of a haplotype A-tagging SNP in *DLG5* which is supported by several other replications of the association between *DLG5* and CD association [34–42]. Results for the two nonsynonymous SNPs in *DLG5* were not reaching significance. We do not consider these SNPs as causative at this point. They are either part of a larger number of putatively causative SNPs not yet discovered or mere additional markers for unknown causative variants. We expect further relevant, hitherto

unknown and rare variants in *DLG5* that may only be detectable by large-scale sequencing of the gene [43]. It should also be noted that the *DLG5* association has not been replicated in each and every population analysed so far [44–46]. Recent studies have proposed genderand/ or age at onset-related associations of *DLG5* [13,34,40–42] that would require exact matching of the study groups to become detectable. Our sample used in this study contained mainly CD patients with disease onset during early adulthood (average age at onset >22 years). This may have contributed to a replication that was weaker than the original description in statistical terms (for review see [47]).

[0124] The targeted replication of the CD association of *ATG16L1*, *IL23R* and *DLG5* also serves to illustrate the highly conservative criteria employed in our study, which may have resulted in an under-appreciation of most initial association findings. Using these criteria, *ATG16L1*/Thr300Ala and *IL23R*/Arg381Gln would not have been included in the follow-up because the p-values of the two variants (0.014 for Thr300Ala and 0.0027 for Arg381Gln) both exceeded the cut-off of 0.0021 (attained by rs3790889 as number 200 of the ranked SNP list). Therefore, future efforts to replicate the major findings of our study should also include those SNPs that yielded a significant p-value in only one of the replication panels.

[0125] The *neural epidermal growth-factor-like* (*nel*) gene was first detected in neural tissue from an embryonic chicken cDNA library, and its human orthologue *NELL1* was later discovered in B-cells [48–50]. The arrangement of the functional domains of the 810 aa protein bears resemblance to thrombospondin-1 (TSP-1) and consists of a thrombospondin N-terminal domain (TSPN) and several von Willebrand factor, type C (VWC), and epidermal growth-factor (EGF) domains [51]. As NELL1 binds to, and is phosphorylated by, PKC-β1 via the EGF domains [52], it has been suggested that this protein belongs to a novel class of cell-signalling ligand molecules critical for growth and development. Re-sequencing and fine mapping revealed several non-synonymous SNPs of which the known Q82R variant and the novel R136S and A153T variants affect the TSPN domain, while R354W is located in a VWC domain (Figure 3) [51]. A153T is close to two highly conserved C-terminal cysteines forming a disulfide bond in the TSPN domain structure of TSP-1 [53] and may cause local conformational changes due to its buried position in the molecule. Generally, the TSPN domain has been shown to serve as a protein-protein interaction module, which binds membrane proteins and proteoglycans and exhibits versatile cell-specific effects on adhesion, migration, and proliferation [54,55]. Since VWC domains occur in numerous proteins of diverse functions and are generally assumed to be involved in protein oligomerization [56], R354W may interfere with *NELL1* trimerization [51]

[0126] Bone development is severely disturbed in transgenic mice, where over-expression of *NELL1* leads to craniosynostis [57] and *NELL1* deficiency manifests in skeletal defects

due to reduced chondro- and osteogenesis [58]. Interestingly, osteopenia and osteoporosis are leading co-morbidities in IBD patients, even without the use of glucocorticoids [59–61]. *PTGER4*, which has been suggested as the causative gene in the *5p13.1* locus, is among the key genes that are down-regulated in *NELL1*-deficient mice [58]. However, no statistical interaction was seen in our study between the *NELL1* and *5p13.1* SNPs.

[0127] The replication criteria used in our study were particularly strict and required a significant p-value in both, the family-based and the case-control association test in two different populations. Other studies used less stringent criteria for the replication of genome-wide association findings, and based their conclusions upon a single independent case control sample only [14,17]. With such criteria, several additional SNPs would have been considered replicated in our study, some of which point towards genes putatively involved in the pathophysiology of IBD. Integrin beta 6 (*ITGB6*), for example, regulates activation of TGF-β [62], a cytokine that has been established as an anti-inflammatory regulator in TNF-related CD pathopysiology [63,64]. Two hits point towards the glutamate pathway, namely glutamate receptor type 8 (*GRM8*) and glutamate receptor, ionotropic, N-methyl-D-aspartate 3A (*GRIN3A*, formerly *PPP3R2*). Normal glutamate metabolism has been found to be important for the maintenance of intestinal function. Finally, SNP rs7868736 is located approximately 100 kb upstream of the *ZNF618* gene encoding a zinc finger protein clearly expressed in human colon.

[0128] In summary, we have successfully performed a systematic genome-wide association scan in Crohn disease that led to the identification of the *NELL1* gene on chromosome *11p15.1* as a novel susceptibility gene for IBD. We confirmed *5p13.1* as a CD-associated locus relating to *PTGER4* that is probably regulated by *NELL1*.

REFERENCES

1. Loftus EV Jr (2004) Clinical epidemiology of inflammatory bowel disease: Incidence, prevalence, and environmental influences. Gastroenterology 126: 1504–1517.

2. Schreiber S, Rosenstiel P, Albrecht M, Hampe J, Krawczak M (2005) Genetics

of Crohn disease, an archetypal inflammatory barrier disease. Nat Rev Genet 6: 376–388.

3. Orholm M, Munkholm P, Langholz E, Nielsen OH, Sorensen TI, et al. (1991) Familial occurrence of inflammatory bowel disease. N Engl J Med 324: 84–88.

4. Kuster W, Pascoe L, Purrmann J, Funk S, Majewski F (1989) The genetics of Crohn disease: complex segregation analysis of a family study with 265 patients with Crohn disease and 5,387 relatives. Am J Med Genet 32: 105–108.

5. Tysk C, Lindberg E, Jarnerot G, 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.

6. Thompson NP, Driscoll R, Pounder RE, Wakefield AJ (1996) Genetics versus environment in inflammatory bowel disease: results of a British twin study. Bmj 312: 95–96.

7. Ogura Y, Bonen DK, Inohara N, Nicolae DL, Chen FF, et al. (2001) A frameshift mutation in NOD2 associated with susceptibility to Crohn's disease. Nature 411: 603–606.

8. Hugot JP, Chamaillard M, Zouali H, Lesage S, Cezard JP, et al. (2001) Association of NOD2 leucine-rich repeat variants with susceptibility to Crohn's disease. Nature 411: 599–603.

9. Hampe J, Cuthbert A, Croucher PJ, Mirza MM, Mascheretti S, et al. (2001) Association between insertion mutation in NOD2 gene and Crohn's disease in German and British populations. Lancet 357: 1925–1928.

10. Peltekova VD, Wintle RF, Rubin LA, Amos CI, Huang Q, et al. (2004) Functional variants of OCTN cation transporter genes are associated with Crohn disease. Nat Genet 36: 471–475.

11. Rioux JD, Daly MJ, Silverberg MS, Lindblad K, Steinhart H, et al. (2001) Genetic variation in the 5q31 cytokine gene cluster confers susceptibility to Crohn disease. Nat Genet 29: 223–228.

12. Stoll M, Corneliussen B, Costello CM, Waetzig GH, Mellgard B, et al. (2004) Genetic variation in DLG5 is associated with inflammatory bowel disease. Nat Genet 36: 476–480.

13. Biank V, Friedrichs F, Babusukumar U, Wang T, Stoll M, et al. (2007) DLG5 R30Q variant is a female-specific protective factor in pediatric onset Crohn's disease. Am J Gastroenterol 102: 391–398.

14. Yamazaki K, McGovern D, Ragoussis J, Paolucci M, Butler H, et al. (2005) Single nucleotide polymorphisms in TNFSF15 confer susceptibility to Crohn's disease. Hum Mol Genet 14: 3499–3506.

15. Hampe J, Franke A, Rosenstiel P, Till A, Teuber M, et al. (2007) A genome-wide association scan of nonsynonymous SNPs identifies a susceptibility variant for Crohn disease in ATG16L1. Nat Genet 39: 207–211.

16. McGovern DP, Hysi P, Ahmad T, van Heel DA, Moffatt MF, et al. (2005) Association between a complex insertion/deletion polymorphism in NOD1 (CARD4) and susceptibility to inflammatory bowel disease. Hum Mol Genet 14: 1245–1250.

17. Duerr RH, Taylor KD, Brant SR, Rioux JD, Silverberg MS, et al. (2006) A genome-wide association study identifies IL23R as an inflammatory bowel disease gene. Science 314: 1461–1463.

18. Libioulle C, Louis E, Hansoul S, Sandor C, Farnir F, et al. (2007) Novel Crohn Disease Locus Identified by Genome-Wide Association Maps to a Gene Desert on 5p13.1 and Modulates Expression of PTGER4. PLoS Genet 3: e58.

19. Cummings JR, Cooney R, Pathan S, Anderson CA, Barrett JC, et al. (2007) Confirmation of the role of ATG16l1 as a Crohn's disease susceptibility gene. Inflamm Bowel Dis.

20. Rioux JD, Xavier RJ, Taylor KD, Silverberg MS, Goyette P, 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.

21. Prescott NJ, Fisher SA, Franke A, Hampe J, Onnie CM, et al. (2007) A Nonsynonymous SNP in ATG16L1 Predisposes to Ileal Crohn's Disease and Is Independent of CARD15 and IBD5. Gastroenterology 132: 1665–1671.

22. Matsuzaki H, Dong S, Loi H, Di X, Liu G, et al. (2004) Genotyping over 100,000 SNPs on a pair of oligonucleotide arrays. Nat Methods 1: 109–111.

23. Fingerlin TE, Boehnke M, Abecasis GR (2004) Increasing the power and efficiency of disease-marker case-control association studies through use of allelesharing information. Am J Hum Genet 74: 432–443.

24. Steffens M, Lamina C, Illig T, Bettecken T, Vogler R, et al. (2006) SNP-based analysis of genetic substructure in the German population. Hum Hered 62: 20–29.

25. Krawczak M, Nikolaus S, von Eberstein H, Croucher PJ, El Mokhtari NE, et al. (2006) PopGen: population-based recruitment of patients and controls for the analysis of complex genotype-phenotype relationships. Community Genet 9: 55–61.

26. Clayton DG, Walker NM, Smyth DJ, Pask R, Cooper JD, et al. (2005) Population structure, differential bias and genomic control in a large-scale, casecontrol association study. Nat Genet 37: 1243–1246.

27. Spielman RS, McGinnis RE, Ewens WJ (1993) Transmission test for linkage disequilibrium: the insulin gene region and insulin-dependent diabetes mellitus (IDDM). Am J Hum Genet 52: 506–516.

28. McIntyre LM, Martin ER, Simonsen KL, Kaplan NL (2000) Circumventing multiple testing: a multilocus Monte Carlo approach to testing for association. Genet Epidemiol 19: 18–29.

29. Fisher SA, Hampe J, Onnie CM, Daly MJ, Curley C, et al. (2006) Direct or indirect association in a complex disease: the role of SLC22A4 and SLC22A5 functional variants in Crohn disease. Hum Mutat 27: 778–785.

30. Parkes M, Barrett JC, Prescott NJ, Tremelling M, Anderson CA, et al. (2007) Sequence variants in the autophagy gene IRGM and multiple other replicating loci contribute to Crohn's disease susceptibility. Nat Genet.

31. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature 447: 661–678.

32. Nicolae DL, Wen X, Voight BF, Cox NJ (2006) Coverage and characteristics of the Affymetrix GeneChip Human Mapping 100K SNP set. PLoS Genet 2: e67.

33. Barrett JC, Cardon LR (2006) Evaluating coverage of genome-wide association studies. Nat Genet 38: 659–662.

34. Lin Z, Poritz LS, Li T, Byrnes KA, Wang Y, et al. (2007) Gender related association of DLG5 variant R30Q with inflammatory bowel disease (IBD) in a familial IBD registry. Gastroenterology 132: A450 (abstract).

35. Daly MJ, Pearce AV, Farwell L, Fisher SA, Latiano A, et al. (2005) Association of DLG5 R30Q variant with inflammatory bowel disease. Eur J Hum Genet 13: 835–839.

36. Newman WG, Gu X, Wintle RF, Liu X, van Oene M, et al. (2006) DLG5 variants contribute to Crohn disease risk in a Canadian population. Hum Mutat 27: 353–358.

37. Yamazaki K, Takazoe M, Tanaka T, Ichimori T, Saito S, et al. (2004) Association analysis of SLC22A4, SLC22A5 and DLG5 in Japanese patients with Crohn disease. J Hum Genet 49: 664–668.

38. Lin Z, Poritz LS, Li T, Byrnes KA, Wang Y, et al. (2007) Genetic variants of DLG5 gene, E514Q, P979L, G1066G, and P1371Q in a familial inflammatory bowel disease (IBD) registry. Gastroenterology 132: A451 (abstract).

39. Weersma RK, Stokkers PC, van Bodegraven AA, van Hogezand RA, Verspaget HW, et al. (2007) A large, nationwide, case-control study for the association of DLG5, OCTN1/2 and CARD15 with Inflammatory Bowel Diseases in the Netherlands. Gastroenterology 132: A445 (abstract).

40. Russell RK, Drummond HE, Nimmo ER, Anderson N, Wilson DC, et al. (2007) The contribution of the DLG5 113A variant in early-onset inflammatory bowel disease. J Pediatr 150: 268–273.

41. de Ridder L, Weersma RK, Dijkstra G, van der Steege G, Benninga MA, et al. (2007) Genetic susceptibility has a more important role in pediatric-onset Crohn's disease than in adult-onset Crohn's disease. Inflamm Bowel Dis advance online publication, DOI: 10.1002/ibd.20171.

42. Friedrichs F, Brescianini S, Annese V, Latiano A, Berger K, et al. (2006) Evidence of transmission ratio distortion of DLG5 R30Q variant in general and implication of an association with Crohn disease in men. Hum Genet 119: 305–311.

43. Romeo S, Pennacchio LA, Fu Y, Boerwinkle E, Tybjaerg-Hansen A, et al. (2007) Population-based resequencing of ANGPTL4 uncovers variations that reduce triglycerides and increase HDL. Nat Genet 39: 513–516.

44. Noble CL, Nimmo ER, Drummond H, Smith L, Arnott ID, et al. (2005) DLG5 variants do not influence susceptibility to inflammatory bowel disease in the Scottish population. Gut 54: 1416–1420.

45. Vermeire S, Pierik M, Hlavaty T, Claessens G, van Schuerbeeck N, et al. (2005) Association of organic cation transporter risk haplotype with perianal penetrating Crohn's disease but not with susceptibility to IBD. Gastroenterology 129: 1845–1853.

46. Torok HP, Glas J, Tonenchi L, Lohse P, Muller-Myhsok B, et al. (2005) Polymorphisms in the DLG5 and OCTN cation transporter genes in Crohn's disease. Gut 54: 1421–1427.

47. Friedrichs F, Stoll M(2006) Role of discs large homolog 5. World J Gastroenterol 12: 3651–3656.

48. Matsuhashi S, Noji S, Koyama E, Myokai F, Ohuchi H, et al. (1996) New gene, nel, encoding a Mr 91 K protein with EGF-like repeats is strongly expressed in neural tissues of early stage chick embryos. Dev Dyn 207: 233–234.

49. Matsuhashi S, Noji S, Koyama E, Myokai F, Ohuchi H, et al. (1995) New gene, nel, encoding a M(r) 93 K protein with EGF-like repeats is strongly expressed in neural tissues of early stage chick embryos. Dev Dyn 203: 212–222.

50. Luce MJ, Burrows PD (1999) The neuronal EGF-related genes NELL1 and NELL2 are expressed in hemopoietic cells and developmentally regulated in the B lineage. Gene 231: 121–126.

51. Kuroda S, Oyasu M, Kawakami M, Kanayama N, Tanizawa K, et al. (1999) Biochemical characterization and expression analysis of neural thrombospondin-1-like proteins NELL1 and NELL2. Biochem Biophys Res Commun 265: 79–86.

52. Kuroda S, Tanizawa K (1999) Involvement of epidermal growth factor-like domain of NELL proteins in the novel protein-protein interaction with protein kinase C. Biochem Biophys Res Commun 265: 752–757.

53. Tan K, Duquette M, Liu JH, Zhang R, Joachimiak A, et al. (2006) The structures of the thrombospondin-1 N-terminal domain and its complex with a synthetic pentameric heparin. Structure 14: 33–42.

54. Bornstein P (1995) Diversity of function is inherent in matricellular proteins: an appraisal of thrombospondin 1. J Cell Biol 130: 503–506.

55. Chen H, Herndon ME, Lawler J (2000) The cell biology of thrombospondin-1. Matrix Biol 19: 597–614.

56. Voorberg J, Fontijn R, Calafat J, Janssen H, van Mourik JA, et al. (1991) Assembly and routing of von Willebrand factor variants: the requirements for disulfide-linked dimerization reside within the carboxy-terminal 151 amino acids. J Cell Biol 113: 195–205.

57. Zhang X, Kuroda S, Carpenter D, Nishimura I, Soo C, et al. (2002) Craniosynostosis in transgenic mice overexpressing Nell-1. J Clin Invest 110: 861–870.

58. Desai J, Shannon ME, Johnson MD, Ruff DW, Hughes LA, et al. (2006) Nell1- deficient mice have reduced expression of extracellular matrix proteins causing cranial and vertebral defects. Hum Mol Genet 15: 1329–1341.

59. Abitbol V, Roux C, Chaussade S, Guillemant S, Kolta S, et al. (1995) Metabolic bone assessment in patients with inflammatory bowel disease. Gastroenterology 108: 417–422.

60. Roux C, Abitbol V, Chaussade S, Kolta S, Guillemant S, et al. (1995) Bone loss in patients with inflammatory bowel disease: a prospective study. Osteoporos Int 5: 156–160.

61. Compston JE, Judd D, Crawley EO, Evans WD, Evans C, et al. (1987) Osteoporosis in patients with inflammatory bowel disease. Gut 28: 410–415.

62. Jenkins RG, Su X, Su G, Scotton CJ, Camerer E, et al. (2006) Ligation of protease-activated receptor 1 enhances alpha(v)beta6 integrin-dependent TGFbeta activation and promotes acute lung injury. J Clin Invest 116: 1606–1614.

63. Monteleone G, Del Vecchio Blanco G, Monteleone I, Fina D, Caruso R, et al. (2005) Post-transcriptional regulation of Smad7 in the gut of patients with inflammatory bowel disease. Gastroenterology 129: 1420–1429.

64. Waetzig GH, Rosenstiel P, Arlt A, Till A, Brautigam K, et al. (2005) Soluble tumor necrosis factor (TNF) receptor-1 induces apoptosis via reverse TNF signaling and autocrine transforming growth factor-beta1. Faseb J 19: 91–93.

65. Croucher PJ, Mascheretti S, Hampe J, Huse K, Frenzel H, et al. (2003) Haplotype structure and association to Crohn's disease of CARD15 mutations in two ethnically divergent populations. Eur J Hum Genet 11: 6–16.

66. Hampe J, Schreiber S, Shaw SH, Lau KF, Bridger S, 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.

67. Lennard-Jones JE (1989) Classification of inflammatory bowel disease. Scand J Gastroenterol Suppl 170: 2–6; discussion 16–19.

68. Truelove SC, Pena AS (1976) Course and prognosis of Crohn's disease. Gut 17: 192–201.

69. Onnie CM, Fisher SA, Pattni R, Sanderson J, Forbes A, et al. (2006) Associations of allelic variants of the multidrug resistance gene (ABCB1 or MDR1) and inflammatory bowel disease and their effects on disease behavior: a case-control and meta-analysis study. Inflamm Bowel Dis 12: 263–271.

70. Hampe J, Wollstein A, Lu T, Frevel HJ, Will M, et al. (2001) An integrated system for high throughput TaqMan based SNP genotyping. Bioinformatics 17: 654–655.

71. Teuber M, Koch W, Manaster C, Waechter S, Hampe J, et al. (2005) Improving quality control and workflow management in high-throughput single-nucleotide polymorphism genotyping environments. Journal of the Association for Laboratory Automation 10: 43–47.

72. Hampe J, Grebe J, Nikolaus S, Solberg C, Croucher PJ, et al. (2002) Association of NOD2 (CARD 15) genotype with clinical course of Crohn's disease: a cohort study. Lancet 359: 1661–1665.

73. Weckx S, Del-Favero J, Rademakers R, Claes L, Cruts M, et al. (2005) novoSNP, a novel computational tool for sequence variation discovery. Genome Res 15: 436–442.

74. Franke A, Wollstein A, Teuber M, Wittig M, Lu T, et al. (2006) GENOMIZER: an integrated analysis system for genome-wide association data. Hum Mutat 27: 583–588.

75. Barrett JC, Fry B, Maller J, Daly MJ (2005) Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics 21: 263–265.

76. Dudbridge F (2003) Pedigree disequilibrium tests for multilocus haplotypes. Genet Epidemiol 25: 115–121.

77. Stephens JC, Schneider JA, Tanguay DA, Choi J, Acharya T, et al. (2001) Haplotype variation and linkage disequilibrium in 313 human genes. Science 293: 489–493.

78. Stephens M, Donnelly P (2003) A comparison of bayesian methods for haplotype reconstruction from population genotype data. Am J Hum Genet 73: 1162–1169.

79. Waetzig GH, Seegert D, Rosenstiel P, Nikolaus S, Schreiber S (2002) p38 mitogen-activated protein kinase is activated and linked to TNF-alpha signaling in inflammatory bowel disease. J Immunol 168: 5342–5351.

CLAIMS

1. A method for determining an irritable bowel disease (IBD) genotype in a patient suspected of having an IBD, or a method of determining a patient's susceptibility to develop an IBD, comprising:

determining the presence or absence of one, or a combination of, single nucleotide polymorphisms (SNPs) in a biological sample from said patient, said SNP(s) being listed in Tables 1-5.

2. The method of claim 1, wherein the SNP(s) are listed in Tables 2-5.

3. The method of claim 1, wherein the presence or absence of two or more SNP(s) selected from rs2076756, rs1992662, rs1992660, rs1793004, rs10521209, and rs2631372 are determined.

4. A method for determining an irritable bowel disease (IBD) genotype in a patient suspected of having an IBD, or a method of determining a patient's susceptibility to develop an IBD, comprising:

determining the presence or absence of a mutation in the gene encoding the nel-like 1 precursor (NELL1) in a biological sample from said patient.

5. The method of claim 4, wherein the mutation is at least one SNP in the NELL1 gene listed in Tables 1, 2, and/or 4.

6. The method of claim 5, wherein the presence or absence of SNP rs1793004 is determined.

7. A method for determining an irritable bowel disease (IBD) genotype in a patient suspected of having an IBD, or a method of determining a patient's susceptibility to develop an IBD, comprising:

determining the presence or absence of a mutation in the 5p13.1 locus in a biological sample from said patient.

8. The method of claim 7, wherein the mutation is at least one SNP listed in Tables 1, 3, and 5.

9. The method of claim 8, wherein the presence or absence of SNP rs1992662 and/or SNP rs1992660 is determined.

10. The method of any one of claims 1 to 9, wherein said patient has a family history of IBD.

11. The method of any one of claims 1 to 10, wherein said patient is suffering from symptoms of an IBD.

12. The method of any one of claims 1 to 11, wherein said patient has a severe IBD phenotype.

13. The method of any one of claims 1 to 12, wherein said IBD is Crohn disease.

14. The method of any one of claims 1 to 12, wherein said IBD is ulcerative colitis.

15. The method of any one of claims 1-14, further comprising, determining the level of expression and/or activity of NELL1 in a biological sample from said patient.

16. The method of any one of claims 1 to 15, further comprising, determining the presence or absence of one or more of the following: a mutation in the CARD15 gene, a mutation in the DLG5 gene, a mutation in the TNFSF15 gene, a mutation in the IL23R gene, and/or a T300A mutation in the ATG16L1 gene.

17. The method of claim 4, further comprising, determining the presence or absence of a mutation in the 5p13.1 locus.

18. The method of claim 17, wherein the presence or absence of SNP rs1992662 and/or SNP rs1992660 is determined in the 5p13.1 locus.

19. The method of any one of claims 1 to 18, wherein the presence or absence of the mutation(s) and/or SNP(s) are determined using a nucleic acid hybridization and/or nucleic acid polymerization assay.

20. The method of claim 19, wherein the presence or absence of the mutation(s) and/or SNP(s) are determined using a gene chip array, TaqMan assay, flap endonuclease assay, or genomic DNA sequencing.

21. The method of any one of claims 1 to 20, wherein the presence or absence of the mutation(s) and/or SNP(s) are determined using a nucleic acid probe specific for each of said mutation(s) and/or SNP(s).

22. A kit or array comprising a set of nucleic acid probes and/or primers specific designed to detect two or more SNPs listed in Tables 1-5, wherein the set of probes and/or primers consists essentially of probes and/or primers related to evaluating an IBD genotype and probes and/or primers related to assay controls.

23. The method of claim 22, wherein the kit or array comprises nucleic acid probes specific for two or more SNP(s) listed in Supplementary Tables 2-5.

24. The method of claim 22, wherein the kit or array comprises nucleic acid probes specific for two or more SNP(s) selected from rs2076756, rs1992662, rs1992660, rs1793004, rs10521209, and/or rs2631372.

25. The method of claim 22, wherein the kit or array comprises nucleic acid probes specific for each of rs2076756, rs1992662, rs1992660, rs1793004, rs10521209, and rs2631372.

26. A NELL1 polypeptide comprising at least one amino acid substitution that is associated with IBD.

27. The polypeptide of claim 26, wherein the polypeptide comprises one or more amino acid substitutions selected from Q82R, R136S, A153T or R354W.

28. A polynucleotide encoding the polypeptide of claim 26 or 27.

29. A host cell harboring the polynucleotide of claim 28.

30. An antibody specific for or raised against the polypeptide of claim 26 or claim 27.

31. A method of treating IBD, comprising administering NELL1 or a functional fragment thereof to a patient in need of treatment.

32. A method for treating IBD, comprising administering an antibody against NELL1 to a patient in need of treatment.

33. A method for identifying an agent for treating IBD, comprising contacting a NELL1 polypeptide with a test agent, and determining a change in the level of NELL1 activity as a result of the test agent.

34. A method for diagnosing sarcoidosis in a subject suspected of having sarcoidosis, or a subject suffering from symptoms of sarcoidosis, comprising, determining the presence or absence of a mutation in NELL1 that is associated with sarcoidosis.

35. The method of claim 34, wherein the SNP is rs951199.

Figure 1
GENI-027/01WO

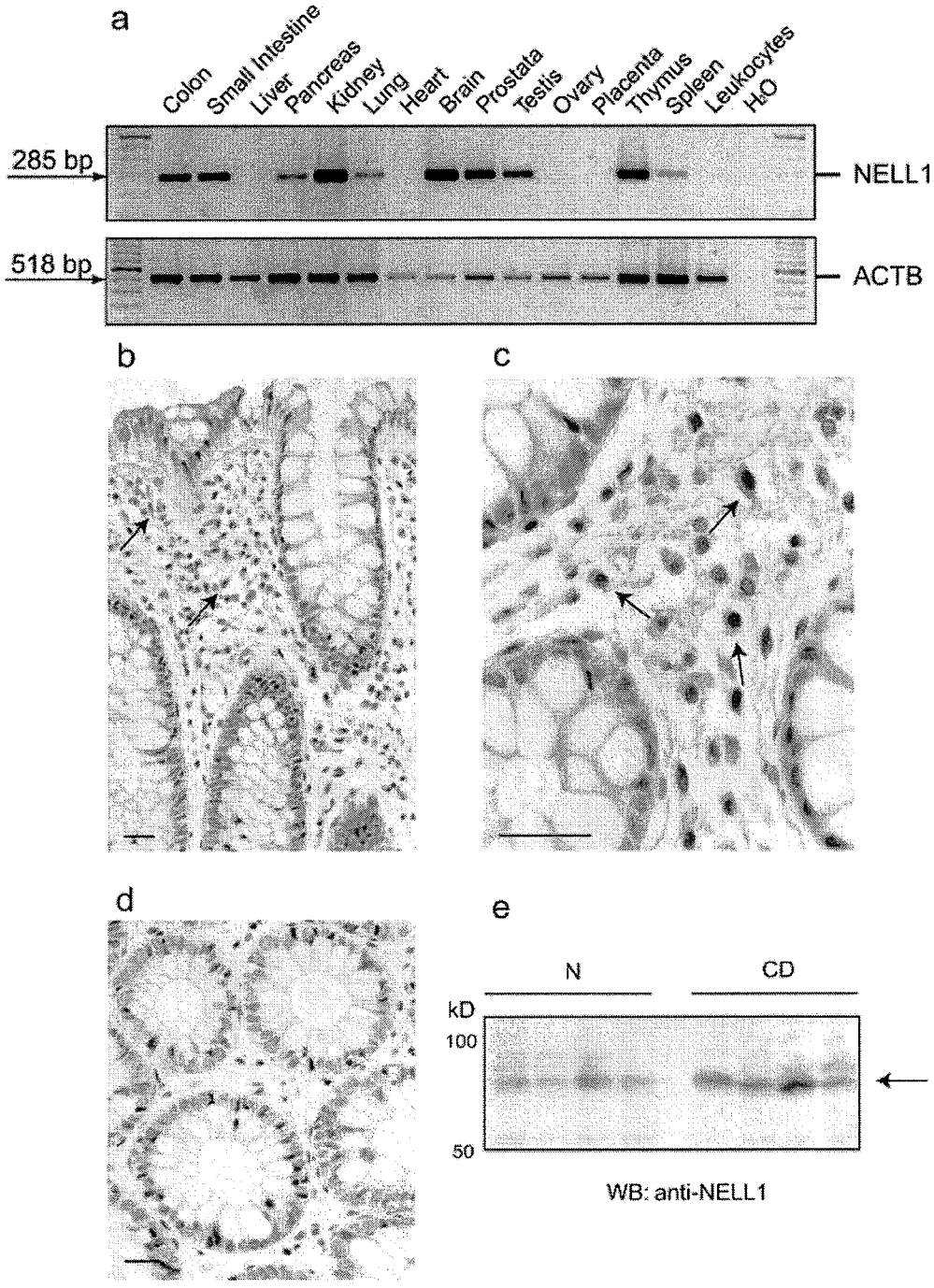


Figure 2A
GENI-027/01WO

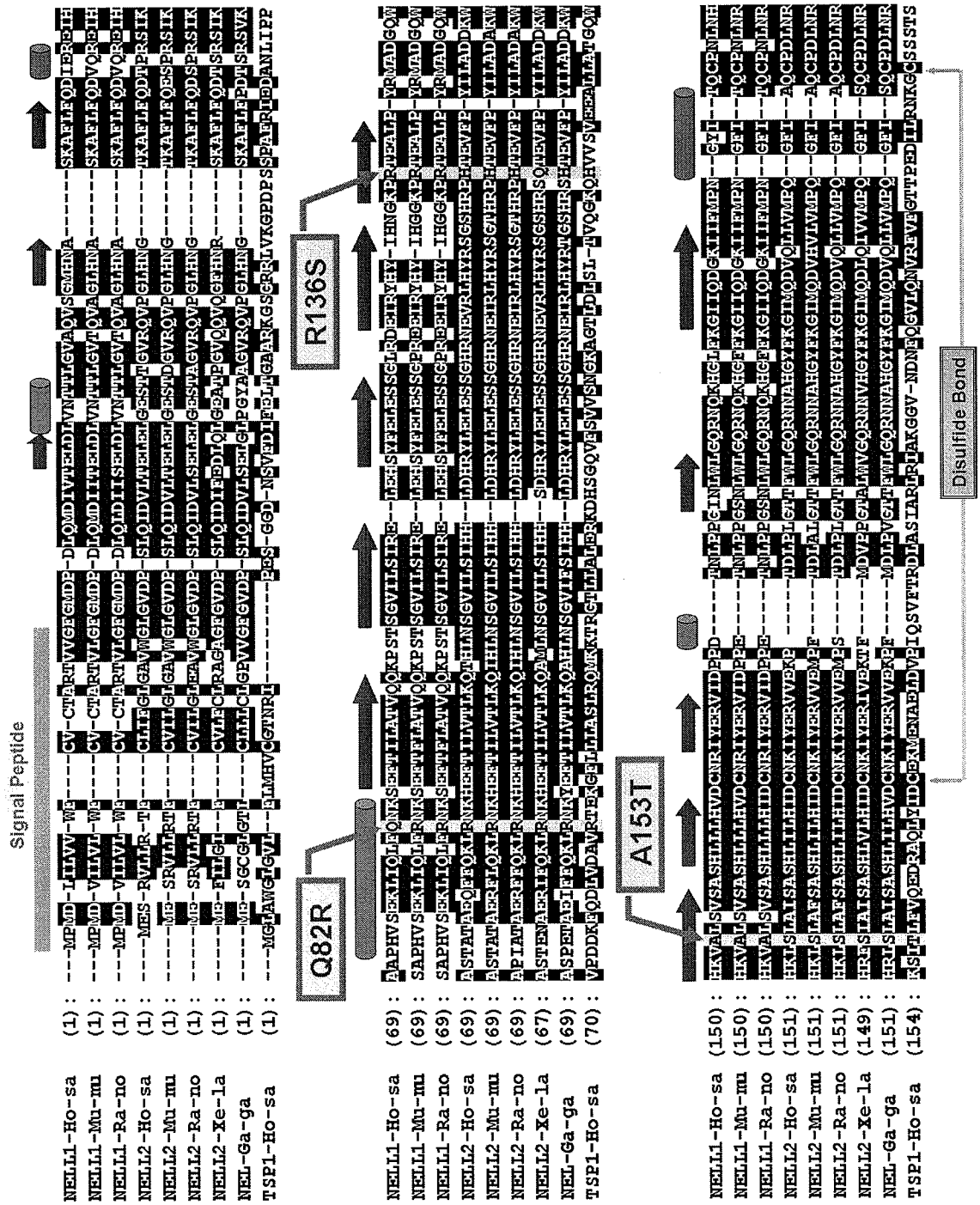


Figure 2B
GENI-027/01WO

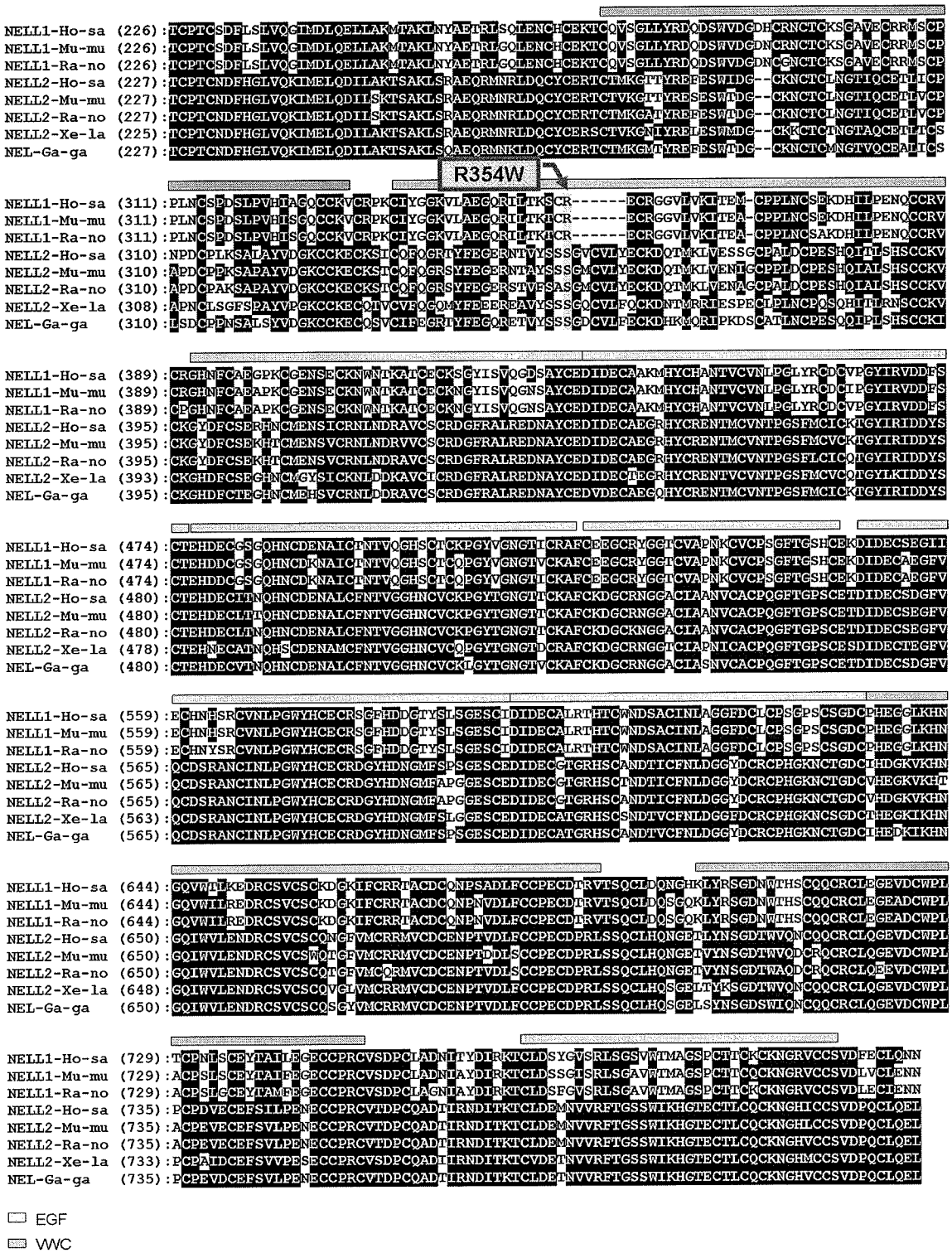


Figure 3
GENI-027/01WO

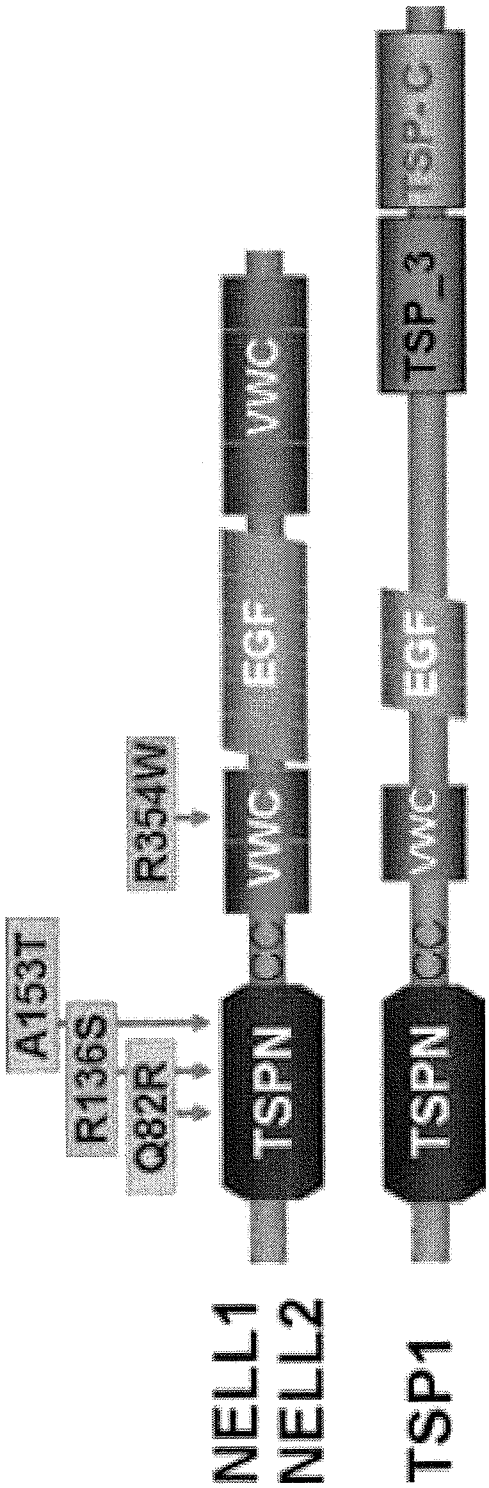


Figure 4
GENI-027/01WO

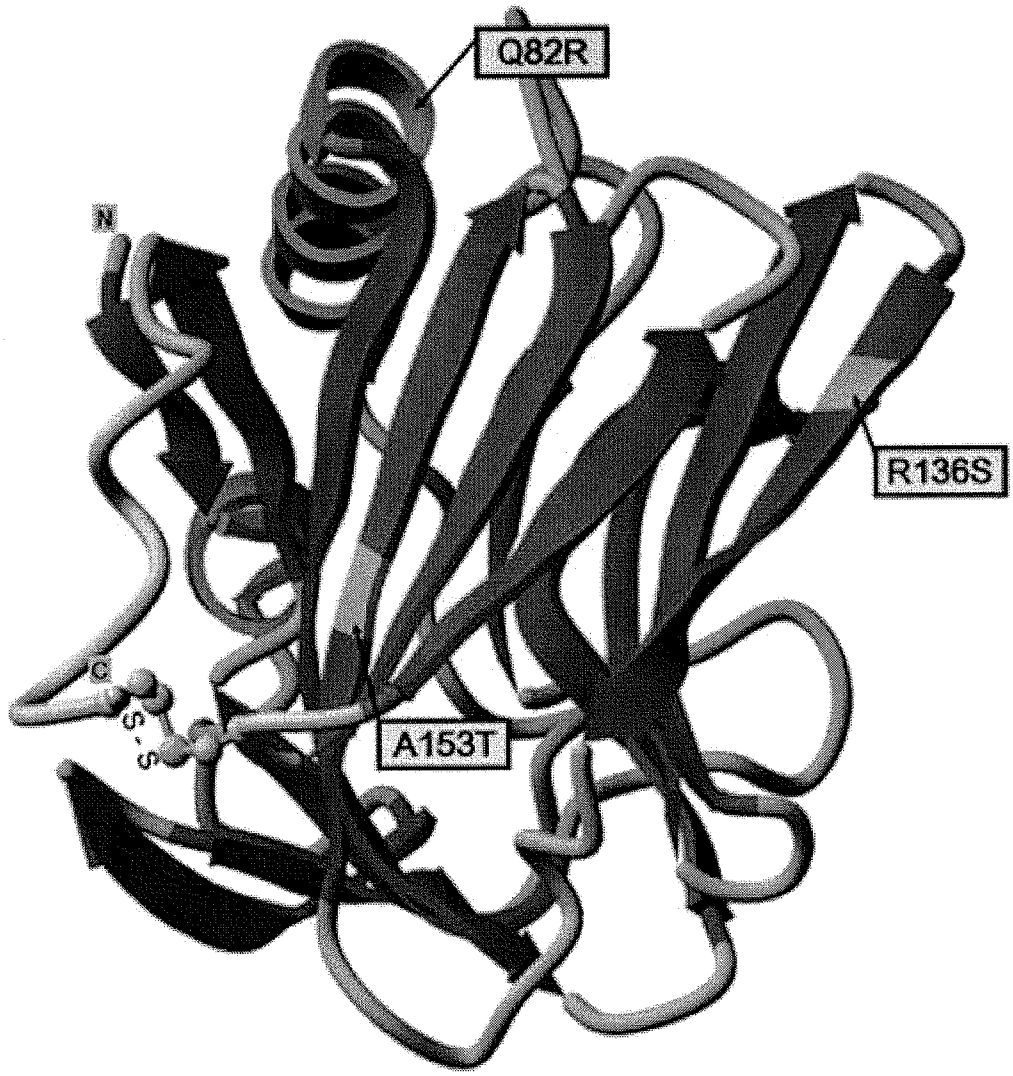


Table 1
GENI-027/01WO

SNP Information						Screening (Panel A)		Replication (Panel B)				
Rank	Affymetrix SNP ID	dbSNP ID	Chr.	Position	Locus	P_{CCA}	P_{CCG}	MAF_{ca}	MAF_{co}	P_{CCA}	P_{CCG}	P_{TDT}
1	*1741384*	*rs2076756*	*16*	*50,534,914*	*CARD15, intron*	*1.93E-13*	*2.04E-12*	*0.27*	*0.41*	*6.80E-20*	*1.39E-21*	*5.90E-08*
2	1683804	rs10513059	3	140,313,327	MRPS22, intron	2.07E-06	1.56E-05	0.12	0.10	0.15	0.14	0.93
3	1700797	rs200354	14	97,365,362	BCL11B, downstream	0.0032	2.22E-06	0.15	0.14	0.59	0.39	1.00
4	1668246	rs2009530	19	53,411,500	CARD8, intron	0.0016	4.92E-06	0.09	0.09	0.73	0.83	0.52
5	1680746	rs10484543	6	28,899,189	ZNF452, upstream	9.36E-06	3.56E-05	0.21	0.19	0.16	0.11	0.066
6	1646788	rs10507063	12	94,769,766	FLJ40089, intron	0.084	1.19E-05	0.11	0.10	0.13	0.060	0.70
7	1664649	rs4978664	9	105,199,766	RAD23B, upstream	1.67E-05	0.00013	0.39	0.38	0.57	0.33	0.66
8	1728972	rs7674505	4	134,995,809	SLC7A11, downstream	2.63E-05	0.00014	0.37	0.36	0.56	0.38	0.13
9	1670981	rs2478126	10	13,663,744	PRPF18, intron	2.70E-05	0.00012	0.06	0.07	0.30	0.29	0.83
10	1716065	rs2660727	4	135,013,676	SLC7A11, downstream	3.95E-05	0.00020	0.37	0.36	0.49	0.21	0.076
11	1741482	rs7324781	13	93,932,754	DZIP1, intron	4.37E-05	0.00030	0.46	0.45	0.46	0.75	0.25
12	1717655	rs6659639	1	14,802,853	BC036877, intron	0.0018	5.17E-05	0.31	0.32	0.68	0.86	0.050
13	1749262	rs1590734	10	117,023,506	ATRNL1, intron	0.0032	5.57E-05	0.47	0.48	0.30	0.11	0.84
14	***1700613***	***rs3859540***	***19***	***60,373,858***	***SYT5, downstream***	***6.18E-05***	***0.00039***	***0.15***	***0.17***	***0.036***	***0.12***	***0.27***
15	***1692451***	***rs6074780***	***20***	***14,701,879***	***BC036403, intron***	***7.12E-05***	***0.00037***	***0.31***	***0.34***	***0.047***	***0.084***	***0.64***
16	1717307	rs10487283	7	102,031,744	FBXL13, intron	0.16	8.60E-05	0.17	0.16	0.37	0.086	0.42
17	1684094	rs10509992	10	117,014,591	ATRNL1, intron	0.0036	8.67E-05	0.47	0.49	0.21	0.097	1.00
18	1747741	rs483952	10	100,048,586	HPSE2, intron	0.0054	8.73E-05	0.19	0.19	0.59	0.71	0.79
19	1752489	rs10494251	1	144,530,971	BCL9, intron	9.83E-05	0.00037	0.06	0.06	0.32	0.23	0.59
20	1666299	rs1559865	18	13,493,887	C18orf1, intron	0.0012	0.00010	0.31	0.34	0.052	0.15	0.30
21	1704045	rs10509357	10	76,373,777	ZNF503, downstream	0.00012	0.00039	0.27	0.28	0.34	0.091	0.52
22	1724833	rs1074742	11	95,528,589	MAML2, intron	0.00012	0.00078	0.39	0.39	0.82	0.98	0.56
23	1722142	rs10517460	4	37,864,263	BC050321, intron	0.011	0.00013	0.11	0.11	0.60	0.85	0.26
24	1750774	rs10493084	1	38,577,447	POU3F1, upstream	0.00014	0.0011	0.03	0.03	0.90	0.38	0.88
25	1728010	rs747066	9	111,078,184	ZFP37, downstream	0.00015	0.00038	0.36	0.36	0.97	0.079	0.86
26	***1717433***	***rs925084***	***4***	***134,989,207***	***SLC7A11, downstream***	***0.00016***	***0.00057***	***0.38***	***0.37***	***0.25***	***0.27***	***0.021***
27	1749044	rs1590736	10	117,023,396	ATRNL1, intron	0.0035	0.00016	0.45	0.48	0.16	0.13	0.87
28	1727572	rs1333261	13	92,604,104	GPC6, intron	0.00018	0.00048	0.26	0.24	0.22	0.46	0.11
29	1724641	rs10501843	11	95,528,273	MAML2, intron	0.00018	0.0013	0.29	0.28	0.48	0.55	0.73
30	1699729	rs2367129	3	193,916,220	LOC151963, intron	0.00019	0.00096	0.28	0.27	0.96	0.91	0.76
31	1690976	rs7092435	10	76,014,914	MYST4, intron	0.00019	0.00043	0.08	0.09	0.79	0.22	0.56
32	1699663	rs605216	13	77,542,856	C13orf10, downstream	0.00020	0.0010	0.08	0.08	0.84	0.11	0.84
33	1666587	rs7700170	4	135,059,812	SLC7A11, downstream	0.00021	0.0011	0.38	0.38	0.62	0.31	0.095
34	***1741147***	***rs4743487***	***9***	***99,916,141***	***CYLC2, upstream***	***0.00021***	***0.0010***	***0.24***	***0.21***	***0.015***	***0.047***	***0.80***
35	***1642811***	***rs2925757***	***2***	***161,303,713***	***ITGB6, upstream***	***0.00021***	***0.0014***	***0.16***	***0.19***	***0.0035***	***0.016***	***1.00***
36	1688004	rs2660693	4	135,078,432	SLC7A11, downstream	0.00022	0.0012	0.27	0.27	0.87	0.98	0.48
37	1698524	rs10501720	11	89,950,836	CHORDC1, upstream	0.00023	0.00075	0.17	0.17	0.82	0.69	0.43
38	1712739	rs853727	4	150,412,639	NR3C2, upstream	0.00024	0.0011	0.26	0.25	0.68	0.85	0.58
39	***1753314***	***rs793515***	***10***	***98,653,056***	***ARHGAP19, intron***	***0.00025***	***0.0010***	***0.31***	***0.30***	***0.65***	***0.89***	***0.0020***
40	1675357	rs10498440	14	50,403,611	GNG2, intron	0.00025	0.00082	0.42	0.44	0.12	0.19	0.50
41	1666766	rs10502288	18	621,435	CLUL1, intron	0.00025	0.0013	0.30	0.29	0.69	0.66	0.076
42	1656805	rs2804200	10	117,014,884	ATRNL1, intron	0.0048	0.00026	0.47	0.49	0.23	0.12	0.79
43	1697779	rs6077888	20	10,687,989	JAG1, upstream	0.00027	0.00082	0.48	0.48	0.57	0.21	0.61
44	1685926	rs7681167	4	76,619,617	DKFZP564O0823, downstream	0.00027	0.0012	0.16	0.16	0.89	0.72	0.35
45	1738725	rs249721	5	115,098,473	CDO1, downstream	0.0061	0.00027	0.19	0.17	0.067	0.20	0.19
46	1708793	rs973542	20	10,687,559	JAG1, upstream	0.00027	0.00082	0.48	0.48	0.57	0.21	0.61
47	1510136	rs1342705	1	144,761,333	GJA5, upstream	0.00027	0.00042	0.21	0.19	0.12	0.28	0.73
48	1716967	rs2928266	5	25,827,114	CDH10, upstream	0.00028	0.0019	0.15	0.13	0.23	0.37	0.39
49	***1673372***	***rs7644124***	***3***	***134,896,774***	***RAB6B, intron***	***0.00029***	***0.00084***	***0.34***	***0.33***	***0.28***	***0.18***	***0.021***
50	1699331	rs3764008	12	20,905,406	SLCO1B3, intron	0.0016	0.00030	0.16	0.16	0.96	1.00	0.47
51	1650842	rs10490876	3	176,472,702	NLGN1, downstream	0.00042	0.00031	0.05	0.06	0.17	0.32	0.39
52	***1740646***	***rs9257453***	***6***	***29,076,919***	***NM_030903, downstream***	***0.00031***	***0.0011***	***0.09***	***0.08***	***0.38***	***0.29***	***0.027***
53	1718827	rs10503282	8	4,626,114	CSMD1, intron	0.00031	0.0016	0.09	0.09	0.98	0.86	0.84
54	1678959	rs3211830	7	79,891,769	CD36, intron	0.00033	0.0019	0.06	0.07	0.34	0.58	0.59
55	1672210	rs10509991	10	117,015,147	ATRNL1, intron	0.0059	0.00037	0.45	0.47	0.30	0.13	0.87
56	1668703	rs7582749	2	37,573,872	QPCT, intron	0.17	0.00037	0.04	0.04	0.91	0.97	0.80
57	1643446	rs7263053	20	47,748,948	SULF2, upstream	0.22	0.00039	0.27	0.29	0.23	0.44	0.46
58	***1662241***	***rs10483112***	***22***	***21,051,522***	***BC030984, intron***	***0.00039***	***0.0020***	***0.03***	***0.02***	***0.025***	***0.024***	***0.37***
59	1712491	rs10496586	2	123,281,730	TSN, downstream	0.00039	0.0018	0.27	0.26	0.27	0.52	0.32
60	1704979	rs10505673	8	138,349,577	KHDRBS3, downstream	0.00071	0.00041	0.21	0.21	0.98	0.69	0.90
61	1713059	rs1845983	4	76,593,158	DKFZP564O0823, downstream	0.00042	0.0018	0.16	0.15	0.48	0.31	0.23
62	1671422	rs1546210	6	92,572,189	MAP3K7, upstream	0.00042	0.0017	0.21	0.21	0.86	0.58	0.19
63	1697996	rs8005131	14	31,581,144	NPAS3, intron	0.00042	0.0031	0.32	0.33	0.44	0.35	0.57
64	***1731332***	***rs7428001***	***3***	***35,500,582***	***PDCD6IP, downstream***	***0.00043***	***0.0013***	***0.44***	***0.44***	***0.84***	***0.97***	***0.012***
65	1740198	rs10508921	10	51,355,311	NM_014114, downstream	0.00043	0.0019	0.08	0.08	0.66	0.82	0.49
66	1748208	rs1559781	4	138,684,038	SLC7A11, downstream	0.00044	0.0028	0.04	0.05	0.65	0.90	0.65
67	1700287	rs4968535	17	59,719,695	BCAS3, intron	0.36	0.00045	0.24	0.25	0.27	0.17	0.60
68	1711405	rs1793483	11	76,701,775	PAK1, upstream	0.00046	0.0016	0.37	0.38	0.41	0.71	0.87
69	1669739	rs10490395	2	56,422,800	KIAA1912, intron	0.0039	0.00047	0.03	0.03	0.70	0.63	0.76
70	*1724612*	*rs1992662*	*5*	*40,439,353*	*PTGER4, upstream*	*0.00047*	*0.0018*	*0.32*	*0.26*	*7.59E-05*	*0.00017*	*0.0013*
71	1751402	rs4323713	1	111,121,947	CHI3L2, downstream	0.00047	0.0033	0.26	0.27	0.60	0.87	0.43
72	***1672987***	***rs6947579***	***7***	***125,087,495***	***GRM8, downstream***	***0.00048***	***0.0022***	***0.29***	***0.33***	***0.0065***	***0.013***	***1.00***
73	1726836	rs812965	3	59,693,634	FHIT, downstream	0.00048	0.0018	0.21	0.20	0.30	0.59	0.30
74	1664846	rs10494257	1	144,761,646	GJA5, upstream	0.00050	0.00080	0.21	0.19	0.16	0.35	0.84
75	*1703018*	*rs1992660*	*5*	*40,460,568*	*PTGER4, upstream*	*0.00050*	*0.0020*	*0.38*	*0.32*	*4.53E-05*	*0.00021*	*0.00050*
76	1650953	rs3922590	10	132,706,022	NM_018461, upstream	0.00050	0.0022	0.12	0.12	0.55	0.75	0.19
77	1691966	rs10498871	6	70,908,433	COL19A1, intron	0.94	0.00051	0.28	0.29	0.39	0.15	0.90
78	1672368	rs3764007	12	20,905,430	SLCO1B3, intron	0.00051	0.0017	0.16	0.16	0.73	0.91	0.34
79	***1689671***	***rs1553575***	***5***	***40,548,433***	***PTGER4, upstream***	***0.00051***	***0.0021***	***0.35***	***0.28***	***1.68E-06***	***6.37E-06***	***0.051***
80	1703646	rs8000906	13	67,561,229	PCDH9, upstream	0.00053	0.0019	0.09	0.09	0.88	0.99	0.83
81	1715786	rs1463446	12	101,488,216	PAH, downstream	0.00053	0.0029	0.39	0.40	0.32	0.60	0.74

Table 1 Continued
GENI-027/01WO

82	1686118	rs10502861	18	41,052,135	SLC14A2, upstream	0.00053	0.0020	0.27	0.29	0.32	0.22	0.86
83	*1663457*	*rs1793004*	*11*	*20,663,238*	*NELL1, intron*	*0.00053*	*0.0020*	*0.28*	*0.24*	*0.025*	*0.052*	*0.045*
84	*1678461*	*rs10487428*	*7*	*125,124,586*	*GRM8, downstream*	*0.00054*	*0.0018*	*0.27*	*0.30*	*0.028*	*0.10*	*0.54*
85	1705772	rs8011773	14	95,847,773	VRK1, downstream	0.00054	0.0041	0.20	0.23	0.057	0.084	0.37
86	*1663676*	*rs10484545*	*6*	*29,342,503*	*OR5V1, downstream*	*0.00054*	*0.0017*	*0.07*	*0.10*	*0.00030*	*2.84E-07*	*0.16*
87	1672449	rs1493517	4	134,930,683	SLC7A11, downstream	0.00054	0.0023	0.37	0.35	0.23	0.48	0.11
88	1755381	rs10517634	4	157,722,979	CTSO, upstream	0.00055	0.00066	0.43	0.42	0.44	0.72	0.21
89	1697863	rs9307696	4	134,945,832	SLC7A11, downstream	0.00057	0.0024	0.37	0.35	0.15	0.35	0.095
90	1670687	rs6673693	1	206,802,809	AY320401, downstream	0.00060	0.0034	0.30	0.32	0.39	0.42	0.73
91	1686876	rs1488927	7	153,794,409	BC035912, downstream	0.0025	0.00061	0.45	0.45	0.86	0.68	0.50
92	1757706	rs1536299	13	22,116,963	MIPEP, intron	0.00061	0.0035	0.23	0.26	0.056	0.13	0.15
93	1665309	rs289386	15	83,311,157	PDE8A, intron	0.00062	0.0026	0.43	0.43	0.87	0.61	0.91
94	*1662678*	*rs723510*	*8*	*78,550,027*	*PXMP3, upstream*	*0.00062*	*0.00069*	*0.36*	*0.32*	*0.037*	*0.093*	*0.069*
95	1724142	rs10514013	18	67,303,699	NM_178506, upstream	0.00078	0.00063	0.06	0.06	0.75	0.72	0.82
96	1743512	rs1413829	9	27,191,968	TEK, intron	0.00064	0.0032	0.28	0.26	0.27	0.45	0.43
97	*1735637*	*rs7795363*	*7*	*31,529,121*	*PDE1C, downstream*	*0.00066*	*0.0026*	*0.16*	*0.15*	*0.34*	*0.027*	*0.71*
98	1728709	rs951308	10	75,962,082	MYST4, intron	0.00067	0.0015	0.09	0.10	0.62	0.56	0.85
99	1664382	rs10486725	7	41,613,489	GLI3, downstream	0.00068	0.0023	0.10	0.10	0.72	0.26	0.93
100	*1694998*	*rs396462*	*8*	*17,680,469*	*FGL1, downstream*	*0.00069*	*0.0013*	*0.21*	*0.22*	*0.46*	*0.46*	*0.030*
101	1643555	rs6122682	20	47,749,305	SULF2, upstream	0.29	0.00070	0.26	0.27	0.69	0.66	0.80
102	*1669289*	*rs10489924*	*1*	*98,921,897*	*AK095544, upstream*	*0.00070*	*0.0024*	*0.10*	*0.10*	*0.65*	*0.57*	*0.026*
103	1734534	rs1859346	5	136,523,790	SPOCK, intron	0.00071	0.0018	0.31	0.30	0.55	0.69	0.25
104	*1658904*	*rs9283536*	*2*	*224,133,045*	*SCG2, downstream*	*0.00071*	*0.0016*	*0.26*	*0.28*	*0.24*	*0.42*	*0.0052*
105	*1653909*	*rs7725523*	*5*	*40,417,724*	*PTGER4, upstream*	*0.00073*	*0.0047*	*0.25*	*0.23*	*0.14*	*0.32*	*0.013*
106	*1721150*	*rs6461599*	*7*	*21,544,775*	*DNAH11, intron*	*0.00074*	*0.0040*	*0.14*	*0.14*	*0.90*	*0.67*	*0.018*
107	1710316	rs1419790	7	34,517,746	BMPER, downstream	0.54	0.00074	0.16	0.17	0.71	0.49	0.25
108	1725734	rs10504252	8	59,267,687	LOC137886, upstream	0.00074	0.0029	0.20	0.19	0.85	0.97	0.072
109	1707295	rs2832238	21	29,469,808	C21orf7, downstream	0.00076	0.0010	0.23	0.23	0.83	0.63	1.00
110	1662590	rs549700	11	88,326,335	GRM5, intron	0.00088	0.00076	0.02	0.02	0.82	0.97	0.73
111	1679422	rs10497212	2	161,167,244	ITGB6, intron	0.00077	0.0036	0.13	0.14	0.41	0.46	0.64
112	1758931	rs4707614	6	91,077,734	MAP3K7, downstream	0.16	0.00078	0.41	0.41	0.96	1.00	0.55
113	1649870	rs10485515	20	14,844,853	BC036403, intron	0.00078	0.0016	0.21	0.22	0.33	0.53	0.85
114	1701880	rs10494248	1	144,615,337	BCL9, downstream	0.00087	0.00079	0.01	0.01	0.67	0.65	0.34
115	*1660459*	*rs10501292*	*11*	*42,960,696*	*API5, upstream*	*0.0098*	*0.00080*	*0.12*	*0.12*	*0.67*	*0.52*	*0.015*
116	*1719121*	*rs951199*	*11*	*20,665,084*	*NELL1, intron*	*0.00080*	*0.0027*	*0.27*	*0.24*	*0.029*	*0.064*	*0.27*
117	*1749234*	*rs10520704*	*15*	*90,450,368*	*SIAT8B, upstream*	*0.00080*	*0.0033*	*0.41*	*0.38*	*0.020*	*0.044*	*0.077*
118	1682582	rs569626	22	26,071,824	CRYBA4, downstream	0.0020	0.00080	0.46	0.44	0.13	0.32	0.33
119	1692328	rs9323605	14	23,606,664	NOVA1, downstream	0.00081	0.0034	0.12	0.11	0.21	0.38	0.62
120	*1678160*	*rs10508621*	*10*	*20,650,445*	*NEBL, downstream*	*0.00082*	*0.0047*	*0.10*	*0.10*	*0.91*	*0.99*	*0.029*
121	1713936	rs10508494	10	15,825,784	C10orf97, intron	0.0037	0.00082	0.24	0.25	0.28	0.33	0.70
122	*1710126*	*rs3118223*	*1*	*48,833,718*	*FLJ11588, intron*	*0.00082*	*0.0033*	*0.40*	*0.37*	*0.026*	*0.084*	*0.47*
123	1647863	rs10500069	7	118,069,959	KCND2, upstream	0.00083	0.0029	0.21	0.21	0.54	0.81	0.90
124	1710468	rs2134066	12	42,000,804	ADAMTS20, downstream	0.00083	0.0029	0.43	0.45	0.24	0.30	0.074
125	*1713324*	*rs4743484*	*9*	*99,859,318*	*PPP3R2, intron*	*0.00083*	*0.0025*	*0.25*	*0.21*	*0.0016*	*0.0069*	*0.90*
126	*1753829*	*rs9304017*	*18*	*8,617,595*	*KIAA0802, upstream*	*0.36*	*0.00084*	*0.14*	*0.16*	*0.082*	*0.22*	*0.042*
127	1656796	rs3096085	5	146,074,928	PPP2R2B, intron	0.0010	0.00085	0.04	0.03	0.44	0.75	0.64
128	1665215	rs10519449	15	21,597,566	NDN, upstream	0.22	0.00086	0.12	0.13	0.79	0.64	0.38
129	*1693727*	*rs10255000*	*7*	*14,785,866*	*DGKB, upstream*	*0.00086*	*0.0038*	*0.15*	*0.13*	*0.025*	*0.088*	*0.32*
130	1748043	rs2852095	18	45,832,038	MYO5B, intron	0.00090	0.0057	0.33	0.33	0.72	0.76	0.82
131	*1668856*	*rs10500037*	*7*	*113,262,290*	*PPP1R3A, upstream*	*0.00091*	*0.0046*	*0.34*	*0.30*	*0.013*	*0.044*	*0.051*
132	1693104	rs1528039	7	110,003,303	IMMP2L, intron	0.00091	0.0032	0.21	0.22	0.75	0.12	0.30
133	1729588	rs770918	1	98,391,860	SNX7, upstream	0.00091	0.0014	0.14	0.13	0.74	0.54	0.44
134	1694761	rs1994893	16	17,879,139	LOC339047, downstream	0.0075	0.00092	0.14	0.14	0.62	0.23	0.33
135	*1705295*	*rs1513287*	*3*	*114,593,109*	*FLJ20174, intron*	*0.00093*	*0.0052*	*0.41*	*0.37*	*0.017*	*0.023*	*1.00*
136	1688914	rs3847621	11	35,328,744	SLC1A2, intron	0.77	0.00093	0.37	0.37	0.60	0.75	0.82
137	1689102	rs2837758	21	40,939,405	DSCAM, intron	0.061	0.00095	0.41	0.40	0.61	0.83	0.55
138	1661156	rs10506296	12	49,374,124	KIAA1463, upstream	0.0011	0.00096	0.03	0.02	0.10	0.18	0.85
139	1691716	rs10503536	8	15,023,600	TUSC3, upstream	0.00096	0.0036	0.20	0.21	0.47	0.72	0.90
140	1657660	rs10492248	12	44,329,880	TMEM16F, downstream	0.00096	0.0042	0.24	0.26	0.11	0.22	0.23
141	1727943	rs10520778	15	93,657,496	FLJ11175, downstream	0.0034	0.00097	0.37	0.36	0.63	0.18	0.34
142	1655695	rs2241330	3	8,832,963	OXTR, upstream	0.00100	0.0041	0.08	0.10	0.12	0.17	0.79
143	1667263	rs7200548	16	5,890,864	A2BP1, upstream	0.0010	0.0035	0.19	0.18	0.78	0.96	0.30
144	1671948	rs10507786	13	79,117,127	BC016673, downstream	0.0010	0.0023	0.22	0.22	0.69	0.92	0.41
145	1748626	rs10500715	11	9,937,371	CMT4B2, intron	0.0011	0.0017	0.45	0.44	0.59	0.17	0.16
146	1675192	rs10512301	9	100,701,999	CYLC2, upstream	0.0011	0.0035	0.06	0.07	0.46	0.60	1.00
147	*1661269*	*rs9304344*	*18*	*43,010,652*	*AY134745, downstream*	*0.0011*	*0.0043*	*0.45*	*0.45*	*0.98*	*0.047*	*0.74*
148	1758667	rs3886870	2	46,017,012	PRKCE, intron	0.0011	0.0051	0.25	0.25	0.68	0.35	0.67
149	1699371	rs1446240	15	57,012,547	RNF111, intron	0.0011	0.0042	0.19	0.17	0.12	0.31	0.51
150	1746102	rs9283944	8	107,290,692	ZFPM2, downstream	0.052	0.0011	0.23	0.25	0.097	0.17	0.096
151	*1751270*	*rs1260126*	*7*	*145,216,060*	*CNTNAP2, intron*	*0.0012*	*0.0056*	*0.23*	*0.19*	*0.011*	*0.015*	*0.17*
152	*1715515*	*rs2123647*	*8*	*95,888,365*	*CCNE2, upstream*	*0.012*	*0.0012*	*0.35*	*0.33*	*0.13*	*0.22*	*0.0088*
153	1667987	rs6501291	17	67,817,577	ABCA10, intron	0.0012	0.0021	0.38	0.37	0.45	0.62	0.42
154	1660643	rs10483431	14	31,671,503	NPAS3, intron	0.11	0.0012	0.14	0.14	0.90	0.95	0.33
155	1695412	rs2173019	5	167,595,866	ODZ2, intron	0.0012	0.0012	0.16	0.16	0.85	0.62	0.66
156	*1674500*	*rs778999*	*5*	*152,936,197*	*GRIA1, intron*	*0.0012*	*0.0032*	*0.34*	*0.31*	*0.016*	*0.037*	*0.14*
157	1645109	rs9302576	16	65,688,397	CDH5, upstream	0.0077	0.0012	0.05	0.04	0.30	0.48	0.89
158	1680152	rs9314133	5	94,600,473	FLJ22344, upstream	0.0012	0.0014	0.26	0.24	0.11	0.11	0.56
159	*1678082*	*rs10521209*	*16*	*50,533,742*	*CARD15, intron*	*0.0013*	*0.0031*	*0.41*	*0.35*	*8.65E-05*	*8.68E-05*	*0.0083*
160	1727370	rs6821328	4	54,133,041	SCFD2, intron	0.59	0.0013	0.47	0.49	0.28	0.46	0.27
161	*1745616*	*rs1161308*	*5*	*99,296,848*	*CHD1, upstream*	*0.65*	*0.0013*	*0.32*	*0.29*	*0.13*	*0.30*	*0.011*
162	*1651987*	*rs10507353*	*13*	*24,600,935*	*CDK8, upstream*	*0.0013*	*0.0064*	*0.50*	*0.49*	*0.75*	*0.92*	*0.019*
163	*1744306*	*rs2631372*	*5*	*131,779,794*	*SLC22A4, downstream*	*0.26*	*0.0013*	*0.33*	*0.29*	*0.0080*	*0.022*	*0.030*

Table 1 Continued
GENI-027/01WO

164	1643625	rs2699750	7	118,043,534	KCND2, upstream	0.0013	0.0021	0.22	0.21	0.67	0.91	0.70
165	1738328	rs4894919	3	106,379,249	ALCAM, upstream	0.0018	0.0013	0.31	0.30	0.38	0.58	0.73
166	1673656	rs10505696	8	139,201,760	COL22A1, downstream	0.0013	0.0024	0.32	0.32	0.75	0.38	0.73
167	1713906	rs1435385	4	37,125,686	KIAA1239, upstream	0.57	0.0013	0.08	0.07	0.40	0.44	0.84
168	1690890	rs10520307	2	82,145,314	SUCLG1, downstream	0.0014	0.0014	0.08	0.08	0.65	0.34	0.68
169	***1720469***	***rs7607642***	***2***	***183,194,719***	***LOC151242, intron***	***0.0022***	***0.0014***	***0.36***	***0.37***	***0.37***	***0.64***	***0.0052***
170	1680241	rs1930551	9	100,720,179	CYLC2, upstream	0.0020	0.0014	0.06	0.07	0.36	0.49	0.50
171	***1699949***	***rs7868736***	***9***	***111,904,228***	***NM_152575, upstream***	***0.0014***	***0.0016***	***0.25***	***0.29***	***0.0025***	***0.0065***	***0.64***
172	1736493	rs1891281	6	14,704,630	CD83, downstream	0.012	0.0014	0.16	0.16	0.76	0.86	0.61
173	***1688588***	***rs1823634***	***11***	***4,405,147***	***SSA1, upstream***	***0.39***	***0.0014***	***0.31***	***0.29***	***0.23***	***0.47***	***0.020***
174	1665706	rs1073298	7	28,107,747	CREB5, upstream	0.20	0.0014	0.16	0.14	0.21	0.43	0.51
175	1702872	rs10487909	7	15,850,165	MEOX2, upstream	0.0017	0.0015	0.03	0.02	0.069	0.14	0.59
176	1714833	rs2823651	21	16,506,279	AF486622, intron	0.0015	0.0067	0.34	0.37	0.062	0.12	0.65
177	1665892	rs10521132	16	25,687,387	HS3ST4, intron	0.036	0.0015	0.24	0.24	0.71	0.62	0.34
178	***1648042***	***rs2964551***	***5***	***114,787,956***	***FEM1C, downstream***	***0.0017***	***0.0015***	***0.02***	***0.02***	***0.97***	***0.64***	***0.016***
179	1722370	rs2385635	2	220,974,685	EPHA4, downstream	0.058	0.0015	0.24	0.22	0.26	0.36	0.57
180	1718706	rs746413	20	41,395,419	PTPRT, intron	0.13	0.0016	0.41	0.40	0.46	0.71	1.00
181	1668809	rs10516960	4	96,618,071	UNC5C, intron	0.0016	0.0026	0.49	0.48	0.60	0.58	0.40
182	1646814	rs1160521	1	191,026,538	B3GALT2, upstream	0.0032	0.0016	0.11	0.10	0.65	0.57	0.21
183	1746058	rs1953915	14	47,150,527	RPS29, downstream	0.0016	0.0024	0.45	0.47	0.25	0.53	0.78
184	1673470	rs2864038	10	106,553,625	SORCS3, intron	0.0018	0.0017	0.02	0.01	0.63	0.63	0.10
185	1758121	rs1420626	12	14,782,776	GUCY2C, upstream	0.0030	0.0017	0.24	0.24	0.94	1.00	0.71
186	1727427	rs491552	6	151,342,046	FTHFSDC1, intron	0.26	0.0017	0.45	0.45	0.68	0.42	0.82
187	1755134	rs10505614	8	134,958,843	ZNF406, downstream	0.018	0.0017	0.29	0.30	0.56	0.84	0.81
188	***1676184***	***rs10515690***	***5***	***152,957,448***	***GRIA1, intron***	***0.0060***	***0.0017***	***0.43***	***0.40***	***0.043***	***0.054***	***0.25***
189	1643487	rs6726390	2	68,051,631	C1D, downstream	0.0017	0.0027	0.11	0.12	0.72	0.91	0.17
190	***1730891***	***rs1695183***	***3***	***23,497,750***	***UBE2E2, intron***	***0.0028***	***0.0018***	***0.22***	***0.19***	***0.10***	***0.22***	***0.023***
191	***1688424***	***rs272867***	***5***	***131,757,273***	***SLC22A4, downstream***	***0.15***	***0.0018***	***0.43***	***0.38***	***0.0034***	***0.0085***	***0.26***
192	***1755639***	***rs830772***	***8***	***76,402,540***	***HNF4G, intron***	***0.012***	***0.0019***	***0.15***	***0.17***	***0.23***	***0.0099***	***0.14***
193	1685410	rs10486718	7	41,400,197	INHBA, downstream	0.0020	0.0019	0.02	0.02	0.12	0.23	0.47
194	***1755609***	***rs747357***	***4***	***9,903,761***	***KIAA1729, downstream***	***0.0019***	***0.011***	***0.43***	***0.46***	***0.085***	***0.010***	***0.92***
195	1736719	rs2025052	14	28,764,583	PRKCM, upstream	0.78	0.0019	0.13	0.14	0.50	0.44	0.87
196	1662317	rs6112307	20	19,284,755	SLC24A3, intron	0.57	0.0020	0.35	0.35	0.93	0.19	0.14
197	***1680706***	***rs1591783***	***1***	***154,422,908***	***ETV3, upstream***	***0.0028***	***0.0021***	***0.47***	***0.43***	***0.024***	***0.074***	***0.38***
198	1657740	rs2023661	16	26,421,292	HS3ST4, downstream	0.0028	0.0021	0.42	0.42	0.84	0.58	0.87
199	1729462	rs7732995	5	106,663,014	EFNA5, downstream	0.0021	0.0077	0.29	0.29	0.83	0.57	0.23
200	1671138	rs3790889	1	81,835,193	LPHN2, intron	0.43	0.0021	0.40	0.40	0.86	0.99	0.74

Table 2
GENI-027/01WO

#	Panel D #	dbSNP ID	Position	Distance [kb]	LD [r^2]	MAF	p_{CCA}	p_{CCG}	p_{TDT}
1	1	rs3740872	20,605,132		0.022	0.63	0.42	0.55	0.87
2	4	rs7122910	20,636,147	31.0	0.000	0.70	0.32	0.60	0.69
3	***5***	***rs1792969***	***20,646,062***	***9.9***	***0.001***	***0.89***	***0.11***	***0.048***	***0.87***
4	***6***	***NELL1_01***	***20,647,024***	***1.0***	***0.007***	***0.99***	***0.57***	***0.56***	***0.034***
5	***7***	***rs1793005***	***20,655,334***	***8.3***	***0.978***	***0.73***	***0.031***	***0.063***	***0.50***
6	***8***	***rs1793004***	***20,655,505***	***0.2***	***0.981***	***0.72***	***0.025***	***0.056***	***0.045***
7	9	rs1793003	20,655,970	0.5	0.983	0.73	0.053	0.096	0.50
8	***10***	***rs951199***	***20,657,351***	***1.4***	***0.083***	***0.73***	***0.029***	***0.071***	***0.27***
9	***11***	***rs870194***	***20,663,881***	***6.5***	***0.509***	***0.79***	***0.28***	***0.25***	***0.048***
10	***12***	***rs1607616***	***20,674,469***	***10.6***	***0.521***	***0.65***	***0.0046***	***0.010***	***0.0010***
11	***13***	***rs1792983***	***20,674,751***	***0.3***	***0.038***	***0.79***	***0.24***	***0.23***	***0.016***
12	14	rs11025705	20,688,454	13.7	0.024	0.86	0.95	0.99	0.20
13	15	rs7952174	20,695,492	7.0	0.004	0.85	0.30	0.54	0.54
14	16	rs908940	20,708,141	12.6	0.004	0.97	0.44	0.15	0.25
15	17	rs7116986	20,711,993	3.9	0.030	0.86	0.93	0.99	0.35
16	18	rs1996623	20,715,992	4.0	0.323	0.85	0.65	0.74	0.27
17	19	rs1519727	20,721,594	5.6	0.390	0.65	0.83	0.49	0.73
18	20	rs1554368	20,732,334	10.7	0.012	0.82	0.82	0.45	0.13
19	21	rs17232778	20,735,857	3.5	0.001	0.92	0.62	0.82	0.28
20	***22***	***rs1429796***	***20,739,579***	***3.7***	***0.005***	***0.98***	***0.0068***	***0.0059***	***0.37***
21	23	rs4922623	20,741,119	1.5	0.363	0.69	0.97	0.98	0.064
22		***rs327025***	***20,746,233***	***5.1***	***0.121***	***0.85***	***0.62***	***0.40***	***0.027***
23	25	rs16906777	20,748,858	2.6	0.025	0.66	0.42	0.73	0.19
24	26	rs327028	20,748,995	0.1	0.237	0.92	0.78	0.84	0.41
25	27	rs1949523	20,750,221	1.2	0.929	0.73	0.100	0.075	0.75
26	***28***	***rs8176785***	***20,761,862***	***11.6***	***0.994***	***0.73***	***0.12***	***0.039***	***0.39***
27	29	rs2280363	20,761,911	0.0	0.353	0.73	0.12	0.075	0.42
28		rs7129413	20,769,224	7.3	0.311	0.87	0.70	0.76	0.48
29	***31***	***rs1158547***	***20,771,723***	***2.5***	***0.539***	***0.67***	***0.34***	***0.37***	***0.023***
30	33	rs11601634	20,791,637	19.9	0.539	0.78	0.35	0.38	0.55
31	***34***	***rs1519735***	***20,791,833***	***0.2***	***0.069***	***0.65***	***0.027***	***0.057***	***0.052***
32	35	rs7109624	20,792,580	0.7	0.035	0.95	0.54	0.11	0.80
33	36	rs7130897	20,796,450	3.9	0.416	0.84	0.75	0.88	0.42
34	37	rs435001	20,800,510	4.1	0.285	0.88	0.25	0.50	0.75
35		rs17298565	20,806,806	6.3	0.004	0.96	0.14	0.16	0.80
36		rs1914984	20,807,913	1.1	0.013	0.92	0.91	0.56	1.00
37	38	rs2680989	20,809,773	1.9	0.511	0.88	0.62	0.49	0.22
38	39	rs919473	20,822,486	12.7	0.755	0.82	0.71	0.44	0.26
39	40	rs7114248	20,824,095	1.6	0.998	0.78	0.79	0.89	0.22
40	41	rs1429799	20,824,886	0.8	0.152	0.79	0.78	0.80	0.30
41		rs1346690	20,827,819	2.9	0.031	0.81	0.80	0.72	0.19
42	43	rs1367002	20,828,374	0.6	0.448	0.89	0.45	0.32	0.93
43	44	rs11025788	20,829,897	1.5	0.406	0.79	0.62	0.70	0.24
44		rs7121400	20,831,014	1.1	0.231	0.90	0.93	0.53	0.13
45		***rs12293297***	***20,832,922***	***1.9***	***0.102***	***0.98***	***0.00095***	***0.0061***	***0.87***
46	45	rs1429794	20,838,013	5.1	0.439	0.82	0.55	0.69	0.29
47		rs6483735	20,844,275	6.3	0.805	0.79	0.10	0.19	0.12
48		rs10766733	20,844,335	0.1	0.212	0.76	0.33	0.25	0.066
49	***46***	***rs919476***	***20,845,201***	***0.9***	***0.012***	***0.93***	***0.19***	***0.19***	***0.042***
50		rs10766735	20,846,648	1.4	0.444	0.83	0.068	0.16	0.50
51		***rs4923128***	***20,848,373***	***1.7***	***0.998***	***0.71***	***0.18***	***0.34***	***0.049***
52	47	rs7109004	20,849,339	1.0	0.167	0.71	0.21	0.43	0.070
53	48	rs1549717	20,867,190	17.9	0.477	0.86	0.80	0.96	0.10
54	50	rs2082080	20,891,024	23.8	0.135	0.75	0.85	0.55	0.14
55	***52***	***rs1880088***	***20,915,761***	***24.7***	***0.141***	***0.75***	***0.14***	***0.029***	***0.49***
56	53	rs8176786	20,915,970	0.2	0.009	0.95	0.63	0.27	0.34
57	54	rs1880084	20,927,365	11.4	0.060	0.73	0.71	0.63	0.074
58	***56***	***rs10833417***	***20,945,048***	***17.7***	***0.326***	***0.74***	***0.018***	***0.045***	***0.49***

Table 2 Continued
GENI-027/01WO

59	57	rs10500884	20,955,364	10.3	0.997	0.88	0.94	0.99	0.74
60	58	rs1400373	20,956,134	0.8	0.015	0.88	0.98	0.94	0.68
61		***rs10500885***	***20,971,636***	***15.5***	***0.000***	***0.90***	***0.049***	***0.022***	***0.87***
62		rs2896623	20,973,973	2.3	0.565	0.84	0.33	0.19	0.36
63	60	rs10500886	20,976,742	2.8	0.922	0.76	0.11	0.29	0.44
64	62	rs11025878	20,998,335	21.6	0.018	0.76	0.31	0.61	0.80
65	63	rs4922728	21,009,790	11.5	0.246	0.81	0.74	0.63	0.55
66	64	rs7937542	21,017,772	8.0	0.308	0.56	0.78	0.25	0.71
67	65	rs10833444	21,029,035	11.3	0.178	0.76	0.49	0.79	0.80
68	66	rs9666195	21,038,467	9.4	0.159	0.65	0.56	0.78	0.56
69	67	rs1543153	21,047,961	9.5	0.200	0.77	0.82	0.96	0.80
70	68	rs7932820	21,056,960	9.0	0.152	0.58	0.94	0.86	0.96
71	69	rs12417046	21,066,312	9.4	0.239	0.81	0.17	0.17	0.84
72	70	rs2403652	21,076,903	10.6	0.273	0.53	0.80	0.29	0.83
73	71	rs7933049	21,087,224	10.3	0.097	0.79	0.58	0.25	0.95
74	72	rs4922753	21,096,666	9.4	0.967	0.71	0.91	0.98	0.51
75	73	rs10766767	21,108,451	11.8	0.143	0.70	0.91	0.99	0.56
76	74	rs6483748	21,137,843	29.4	0.642	0.70	0.71	0.37	0.44
77	75	rs4475918	21,147,706	9.9	0.025	0.60	0.35	0.39	0.60
78	77	rs1453983	21,166,668	19.0	0.647	0.75	0.95	0.81	0.53
79	78	rs1454003	21,175,735	9.1	0.149	0.83	0.21	0.46	0.32
80	79	rs1823843	21,185,111	9.4	0.251	0.97	0.85	0.36	0.78
81	80	rs1945327	21,190,821	5.7	0.028	0.90	0.61	0.83	0.80
82	81	rs1453988	21,201,279	10.5	0.983	0.69	0.47	0.75	0.32
83	82	rs1670638	21,207,776	6.5	0.033	0.68	0.69	0.88	0.34
84	83	rs1670640	21,218,026	10.3	0.026	0.54	0.74	0.15	0.64
85	84	rs1454008	21,230,114	12.1	0.670	0.59	0.53	0.74	0.57
86	85	rs1791822	21,240,131	10.0	0.026	0.54	0.65	0.12	0.87
87	86	rs10500901	21,240,719	0.6	0.028	0.97	0.84	0.98	0.66
88	87	rs1453990	21,249,027	8.3	0.484	0.53	0.57	0.48	0.91
89	88	rs716577	21,262,140	13.1	0.957	0.63	0.46	0.62	0.42
90	89	rs6483756	21,267,196	5.1	0.752	0.62	0.24	0.46	0.41
91	***90***	***rs10833498***	***21,277,881***	***10.7***	***0.621***	***0.68***	***0.32***	***0.60***	***0.023***
92	91	rs4335544	21,286,758	8.9	0.756	0.59	0.83	0.86	0.52
93	92	rs1349818	21,292,747	6.0	0.001	0.51	0.28	0.37	0.74
94	93	rs4399327	21,302,941	10.2	0.001	0.52	0.52	0.093	1.00
95	94	rs2187522	21,313,688	10.7	0.160	0.51	0.91	0.73	0.25
96	95	rs11026036	21,323,317	9.6	0.132	0.64	0.44	0.73	0.87
97	96	rs7126959	21,333,727	10.4	0.015	0.71	0.32	0.59	0.95
98	97	rs1945404	21,343,840	10.1	0.031	0.63	0.81	0.13	1.00
99	98	rs4151056	21,349,063	5.2	0.006	0.95	0.068	0.18	0.41
100	99	rs10833520	21,352,936	3.9	0.009	0.80	0.79	0.47	0.18
101	100	rs1945443	21,365,895	13.0	0.009	0.66	0.58	0.25	0.61
102	***101***	***rs4539321***	***21,375,322***	***9.4***	***0.023***	***0.54***	***0.45***	***0.033***	***0.66***
103	102	rs11026072	21,385,901	10.6	0.169	0.67	0.47	0.58	0.066
104	103	rs7943922	21,394,268	8.4	0.081	0.50	0.085	0.22	0.20
105	104	rs11026079	21,406,963	12.7	0.065	0.75	0.59	0.50	1.00
106	105	rs7110569	21,418,064	11.1	0.030	0.83	0.88	0.97	0.56
107	106	rs1945408	21,428,394	10.3	0.125	0.86	0.88	0.62	0.36
108	107	rs7945802	21,438,700	10.3	0.445	0.57	0.66	0.48	0.83
109	109	rs10766821	21,458,271	19.6	0.097	0.62	0.72	0.91	0.54
110	110	rs7116826	21,472,159	13.9	0.027	0.85	0.44	0.60	0.39
111	111	rs10219188	21,479,105	6.9	0.465	0.53	0.48	0.74	0.87
112	112	rs6483774	21,491,019	11.9	0.244	0.51	0.55	0.81	0.96
113	113	rs10766829	21,499,146	8.1	0.836	0.59	0.45	0.41	0.91
114	116	rs7926887	21,523,755	24.6	0.599	0.61	0.53	0.54	0.96
115	117	rs4319515	21,534,943	11.2	0.202	0.51	0.14	0.34	0.96
116	118	rs8176789	21,538,381	3.4	0.426	0.83	0.86	0.50	0.94
117	120	rs4922850	21,557,214	18.8		0.93	0.30	0.097	0.58

Table 3
GENI-027/01WO

#	Panel D #	dbSNP ID	Position	Distance [kb]	LD [r^2]	MAF	p_{CCA}	p_{CCG}	p_{TDT}
1		*rs189814*	*40,292,083*		*0.400*	*0.488*	*0.25*	*0.21*	*0.041*
2	*19*	*rs1445011*	*40,315,959*	*23.9*	*0.000*	*0.279*	*0.052*	*0.078*	*0.011*
3		*rs1445002*	*40,355,634*	*39.7*	*0.048*	*0.138*	*3.94E-05*	*0.00021*	*0.042*
4	*41*	*rs443583*	*40,368,840*	*13.2*	*0.250*	*0.206*	*0.061*	*0.16*	*0.022*
5	44	rs12518245	40,381,478	12.6	0.181	0.060	0.50	0.70	0.91
6	*56*	*rs7725523*	*40,407,980*	*26.5*	*0.359*	*0.254*	*0.14*	*0.32*	*0.013*
7	*67*	*rs1992662*	*40,429,609*	*21.6*	*0.741*	*0.321*	*7.59E-05*	*0.00017*	*0.0013*
8	*78*	*rs1992660*	*40,450,824*	*21.2*	*0.474*	*0.380*	*4.53E-05*	*0.00021*	*0.00050*
9	*100*	*rs4957300*	*40,499,496*	*48.7*	*0.263*	*0.330*	*0.0033*	*0.0061*	*0.00020*
10		*rs4422570*	*40,518,098*	*18.6*	*0.056*	*0.099*	*0.83*	*0.75*	*0.030*
11	*107*	*rs7725052*	*40,523,027*	*4.9*	*0.572*	*0.442*	*3.24E-06*	*1.95E-05*	*0.0017*
12	*113*	*rs1553575*	*40,538,689*	*15.7*	*0.482*	*0.348*	*1.68E-06*	*6.37E-06*	*0.033*
13		*rs7718309*	*40,564,656*	*26.0*	*0.912*	*0.186*	*0.023*	*0.013*	*0.085*
14	132	rs7713972	40,588,231	23.6	0.553	0.161	0.26	0.15	0.14
15	*133*	*rs6451525*	*40,590,026*	*1.8*	*0.498*	*0.296*	*1.64E-05*	*4.49E-05*	*0.039*
16	*136*	*rs6864749*	*40,604,350*	*14.3*	*0.414*	*0.185*	*0.029*	*0.014*	*0.12*
17	140	rs6451529	40,611,106	6.8	0.712	0.079	0.66	0.68	0.27
18	*151*	*rs4409138*	*40,639,362*	*28.3*	*1.000*	*0.118*	*0.047*	*0.14*	*0.35*
19	156	rs924967	40,650,879	11.5	0.014	0.117	0.055	0.16	0.40
20		rs6877753	40,677,032	26.2	0.001	0.114	0.50	0.56	0.52
21	184	rs4546432	40,718,307	41.3	0.742	0.378	0.43	0.20	0.37
22	186	rs6451535	40,723,788	5.5	0.349	0.323	0.98	0.27	0.22
23		rs10053664	40,821,568	97.8		0.430	0.43	0.48	0.46

Table 4
GENI-027/01WO

#	dbSNP ID	Position	Distance [kb]	LD [r^2]	MAF	T:U	P_{TDT}
1	rs3740872	20,605,132		0.114	0.420	189:162	0.15
2	rs2000959	20,615,255	10.1	0.435	0.324	160:143	0.33
3	***rs10766710***	***20,624,679***	***9.4***	***0.032***	***0.499***	***195:156***	***0.037***
4	rs7122910	20,636,147	11.5	0.000	0.294	145:138	0.68
5	rs1792969	20,646,062	9.9	0.000	0.128	79:66	0.28
6	NELL1_01	20,647,024	1.0	0.012	0.004	3:2	0.65
7	***rs1793005***	***20,655,334***	***8.3***	***0.994***	***0.263***	***156:117***	***0.018***
8	***rs1793004***	***20,655,505***	***0.2***	***0.966***	***0.263***	***140:107***	***0.036***
9	***rs1793003***	***20,655,970***	***0.5***	***0.953***	***0.261***	***157:119***	***0.022***
10	***rs951199***	***20,657,351***	***1.4***	***0.089***	***0.264***	***151:117***	***0.038***
11	rs870194	20,663,881	6.5	0.505	0.199	116:102	0.34
12	rs1607616	20,674,469	10.6	0.515	0.327	152:140	0.48
13	rs1792983	20,674,751	0.3	0.042	0.200	117:102	0.31
14	rs11025705	20,688,454	13.7	0.029	0.146	87:75	0.35
15	***rs7952174***	***20,695,492***	***7.0***	***0.003***	***0.145***	***98:71***	***0.038***
16	rs908940	20,708,141	12.6	0.003	0.020	14:11	0.55
17	rs7116986	20,711,993	3.9	0.031	0.133	92:68	0.058
18	rs1996623	20,715,992	4.0	0.338	0.167	95:83	0.37
19	rs1519727	20,721,594	5.6	0.322	0.365	162:146	0.36
20	rs1554368	20,732,334	10.7	0.015	0.155	102:77	0.062
21	rs17232778	20,735,857	3.5	0.003	0.077	47:39	0.39
22	rs1429796	20,739,579	3.7	0.017	0.034	24:17	0.27
23	rs4922623	20,741,119	1.5	0.496	0.328	144:137	0.68
24	rs327036	20,742,093	1.0	0.411	0.194	118:99	0.20
25	rs16906777	20,748,858	6.8	0.034	0.343	141:141	1.00
26	rs327028	20,748,995	0.1	0.255	0.093	50:47	0.76
27	***rs1949523***	***20,750,221***	***1.2***	***0.930***	***0.255***	***149:110***	***0.015***
28	rs8176785	20,761,862	11.6	1.000	0.251	136:109	0.085
29	rs2280363	20,761,911	0.0	0.016	0.252	135:108	0.083
30	rs17298537	20,765,554	3.6	0.236	0.089	61:48	0.21
31	***rs1158547***	***20,771,723***	***6.2***	***0.931***	***0.288***	***151:118***	***0.044***
32	***rs4923055***	***20,779,494***	***7.8***	***0.555***	***0.273***	***148:114***	***0.036***
33	***rs11601634***	***20,791,637***	***12.1***	***0.454***	***0.173***	***100:73***	***0.040***
34	***rs1519735***	***20,791,833***	***0.2***	***0.081***	***0.314***	***160:125***	***0.038***
35	rs7109624	20,792,580	0.7	0.012	0.047	29:21	0.26
36	rs7130897	20,796,450	3.9	0.300	0.145	85:70	0.23
37	rs435001	20,800,510	4.1	0.018	0.122	76:67	0.45
38	rs2680989	20,809,773	9.3	0.450	0.112	69:63	0.60
39	rs919473	20,822,486	12.7	0.749	0.179	99:88	0.42
40	rs7114248	20,824,095	1.6	1.000	0.224	113:109	0.79
41	rs1429799	20,824,886	0.8	0.001	0.224	113:109	0.79
42	NELL1_03	20,825,826	0.9	0.001	0.005	4:3	0.71
43	rs1367002	20,828,374	2.5	0.495	0.126	71:67	0.73
44	rs11025788	20,829,897	1.5	0.001	0.223	119:111	0.60
45	rs1429794	20,838,013	8.1	0.012	0.186	108:101	0.63
46	rs919476	20,845,201	7.2	0.151	0.049	33:26	0.36
47	rs7109004	20,849,339	4.1	0.251	0.256	130:127	0.85
48	rs1549717	20,867,190	17.9	0.586	0.139	88:81	0.59
49	rs4296038	20,878,420	11.2	0.883	0.213	118:111	0.64
50	rs2082080	20,891,024	12.6	0.078	0.224	119:115	0.79
51	rs2293241	20,905,691	14.7	0.889	0.278	136:130	0.71
52	rs1880088	20,915,761	10.1	0.153	0.256	130:126	0.80
53	rs8176786	20,915,970	0.2	0.013	0.049	37:30	0.39
54	rs1880084	20,927,365	11.4	0.039	0.302	138:135	0.86
55	rs952696	20,935,188	7.8	0.523	0.197	110:103	0.63
56	rs10833417	20,945,048	9.9	0.401	0.270	142:141	0.95
57	rs10500884	20,955,364	10.3	1.000	0.131	72:72	1.00
58	rs1400373	20,956,134	0.8	0.041	0.131	73:73	1.00
59	rs10766756	20,966,312	10.2	0.081	0.131	82:78	0.75
60	rs10500886	20,976,742	10.4	0.544	0.246	128:117	0.48
61	rs12791452	20,986,789	10.0	0.476	0.151	91:80	0.40
62	rs11025878	20,998,335	11.5	0.039	0.260	122:122	1.00
63	rs4922728	21,009,790	11.5	0.207	0.190	107:106	0.95
64	rs7937542	21,017,772	8.0	0.331	0.455	181:168	0.49
65	rs10833444	21,029,035	11.3	0.156	0.268	128:120	0.61
66	rs9666195	21,038,467	9.4	0.188	0.361	156:147	0.61
67	rs1543153	21,047,961	9.5	0.214	0.252	124:117	0.65
68	rs7932820	21,056,960	9.0	0.156	0.395	165:161	0.82
69	rs12417046	21,066,312	9.4	0.254	0.192	107:102	0.73
70	rs2403652	21,076,903	10.6	0.284	0.481	176:162	0.45
71	rs7933049	21,087,224	10.3	0.098	0.218	119:114	0.74
72	rs4922753	21,096,666	9.4	0.959	0.262	132:128	0.80
73	rs10766767	21,108,451	11.8	0.120	0.257	130:126	0.80
74	rs6483748	21,137,843	29.4	0.620	0.285	142:120	0.17
75	rs4475918	21,147,706	9.9	0.543	0.391	163:148	0.40
76	rs4923403	21,156,380	8.7	0.125	0.293	142:121	0.20
77	rs1453983	21,166,668	10.3	0.627	0.261	145:118	0.096
78	rs1454003	21,175,735	9.1	0.075	0.187	113:101	0.41
79	rs1823843	21,185,111	9.4	0.157	0.017	15:9	0.22
80	rs1945327	21,190,821	5.7	0.023	0.087	60:43	0.094
81	rs1453988	21,201,279	10.5	0.985	0.313	149:149	1.00
82	rs1670638	21,207,776	6.5	0.041	0.313	149:148	0.95
83	rs1670640	21,218,026	10.3	0.031	0.475	166:166	1.00
84	rs1454008	21,230,114	12.1	0.699	0.418	159:150	0.61
85	rs1791822	21,240,131	10.0	0.025	0.455	161:159	0.91
86	rs10500901	21,240,719	0.6	0.025	0.029	25:20	0.46
87	rs1453990	21,249,027	8.3	0.466	0.473	160:155	0.78
88	rs716577	21,262,140	13.1	0.922	0.369	166:159	0.70
89	rs6483756	21,267,196	5.1	0.766	0.380	166:158	0.66
90	rs10833498	21,277,881	10.7	0.607	0.325	161:151	0.57
91	rs4335544	21,286,758	8.9	0.771	0.413	176:165	0.55
92	rs1349818	21,292,747	6.0	0.000	0.470	178:161	0.36
93	rs4399327	21,302,941	10.2	0.026	0.500	182:165	0.36
94	rs2187522	21,313,688	10.7	0.121	0.484	193:171	0.25
95	rs11026036	21,323,317	9.6	0.149	0.352	173:154	0.29
96	rs7126959	21,333,727	10.4	0.017	0.279	154:130	0.15
97	rs1945404	21,343,840	10.1	0.030	0.343	158:147	0.53
98	rs4151056	21,349,063	5.2	0.007	0.053	31:29	0.80
99	rs10833520	21,352,936	3.9	0.008	0.212	125:120	0.75
100	rs1945443	21,365,895	13.0	0.007	0.380	154:143	0.52
101	rs4539321	21,375,322	9.4	0.019	0.473	178:160	0.33
102	rs11026072	21,385,901	10.6	0.124	0.341	165:155	0.58
103	rs7943922	21,394,268	8.4	0.132	0.481	177:171	0.75
104	rs11026079	21,406,963	12.7	0.054	0.286	145:141	0.81
105	rs7110569	21,418,064	11.1	0.026	0.131	77:68	0.45
106	rs1945408	21,428,394	10.3	0.139	0.166	90:86	0.76
107	rs7945802	21,438,700	10.3	0.132	0.420	172:159	0.47
108	rs4343021	21,446,511	7.8	0.113	0.159	89:85	0.76
109	rs10766821	21,458,271	11.8	0.096	0.374	159:150	0.61
110	rs7116826	21,472,159	13.9	0.040	0.174	98:94	0.77
111	rs10219188	21,479,105	6.9	0.444	0.476	185:175	0.60
112	rs6483774	21,491,019	11.9	0.227	0.476	189:171	0.34
113	rs10766829	21,499,146	8.1	0.058	0.451	171:166	0.79
114	rs6483779	21,509,362	10.2	0.067	0.063	48:34	0.12
115	rs7927068	21,516,670	7.3	0.899	0.451	168:163	0.78
116	rs7926887	21,523,755	7.1	0.527	0.435	178:168	0.59
117	rs4319515	21,534,943	11.2	0.202	0.430	180:148	0.077
118	***rs8176789***	***21,538,381***	***3.4***	***0.206***	***0.212***	***124:92***	***0.030***
119	rs4320947	21,544,279	5.9	0.100	0.449	177:147	0.096
120	rs4922850	21,557,214	12.9		0.110	72:60	0.30

Table 5
GENI-027/01WO

#	dbSNP ID	Position	Distance [kb]	LD [r²]	MAF	T:U	PTDT
1	rs350031	40,223,211		0.004	0.318	128:118	0.52
2	rs16869727	40,229,955	6.7	0.013	0.131	69:60	0.43
3	rs353325	40,246,690	16.7	0.008	0.177	80:74	0.63
4	rs183046	40,269,981	23.3	0.021	0.256	99:95	0.77
5	rs17224401	40,275,446	5.5	0.001	0.303	119:109	0.51
6	rs350052	40,290,787	15.3	0.002	0.196	112:104	0.59
7	rs238772	40,292,388	1.6	0.000	0.001	1:0	0.32
8	rs2371680	40,294,274	1.9	0.000	0.289	163:131	0.062
9	rs17823523	40,294,659	0.4	0.000	0.446	127:118	0.57
10	rs348609	40,296,686	2.0	0.256	0.193	111:105	0.68
11	rs6881572	40,299,361	2.7	0.434	0.483	188:170	0.34
12	rs1348612	40,301,661	2.3	0.002	0.290	163:131	0.062
13	***rs1445010***	***40,304,133***	***2.5***	***0.002***	***0.155***	***86:55***	***0.0090***
14	rs755989	40,307,432	3.3	0.433	0.289	163:131	0.062
15	rs2120857	40,310,215	2.8	0.434	0.484	190:170	0.29
16	rs971212	40,311,451	1.2	1.000	0.290	163:131	0.062
17	rs4957262	40,313,082	1.6	1.000	0.289	163:130	0.054
18	rs2860001	40,314,606	1.5	0.009	0.289	163:131	0.062
19	rs1445011	40,315,959	1.4	0.009	0.000	147:119	0.086
20	rs1445012	40,317,863	1.9	1.000	0.289	162:131	0.070
21	rs1373696	40,319,579	1.7	1.000	0.289	162:131	0.070
22	rs17225890	40,321,140	1.6	0.003	0.289	162:131	0.070
23	rs348620	40,322,578	1.4	0.029	0.109	63:45	0.083
24	rs4245975	40,324,635	2.1	0.996	0.182	107:89	0.20
25	rs12518811	40,327,354	2.7	0.010	0.181	107:88	0.17
26	rs348584	40,329,072	1.7	0.003	0.453	139:135	0.81
27	rs348589	40,331,377	2.3	0.350	0.153	90:79	0.40
28	rs1564269	40,338,324	6.9	0.350	0.337	154:151	0.86
29	rs348571	40,339,742	1.4	0.394	0.153	90:79	0.40
30	rs180900	40,342,058	2.3	0.004	0.255	108:94	0.32
31	rs10512731	40,344,547	2.5	0.050	0.224	124:110	0.36
32	rs348560	40,348,617	4.1	0.005	0.153	90:79	0.40
33	rs6871840	40,350,199	1.6	0.006	0.029	22:14	0.18
34	rs4957127	40,351,767	1.6	0.257	0.161	77:65	0.31
35	rs348601	40,355,763	4.0	0.400	0.386	179:157	0.23
36	rs348599	40,356,915	1.2	0.394	0.202	120:102	0.23
37	rs348595	40,359,471	2.6	0.062	0.385	176:158	0.32
38	rs10043093	40,359,977	0.5	0.017	0.249	105:98	0.62
39	rs348593	40,360,289	0.3	0.993	0.197	118:102	0.28
40	rs2034185	40,361,415	1.1	1.000	0.196	117:100	0.25
41	rs443583	40,368,840	7.4	0.993	0.000	118:101	0.25
42	rs394213	40,369,508	0.7	0.990	0.197	118:102	0.28
43	rs7726744	40,379,033	9.5	0.291	0.196	118:101	0.25
44	rs12518245	40,381,478	2.4	0.010	0.000	51:43	0.41
45	***rs16869833***	***40,383,331***	***1.9***	***0.097***	***0.241***	***113:73***	***0.0034***
46	rs4286720	40,385,373	2.0	0.087	0.175	99:94	0.72
47	rs7725639	40,391,593	6.2	0.995	0.293	153:147	0.73
48	rs6451489	40,393,420	1.8	0.090	0.293	154:147	0.69
49	rs12517783	40,395,897	2.5	1.000	0.182	100:98	0.89
50	rs4957278	40,397,577	1.7	1.000	0.181	98:97	0.94
51	rs12514900	40,399,633	2.1	0.017	0.182	100:98	0.89
52	rs6451491	40,401,673	2.0	0.017	0.071	49:42	0.46
53	rs12514679	40,402,560	0.9	0.186	0.182	98:98	1.00
54	rs10472332	40,404,130	1.6	0.001	0.456	191:176	0.43
55	rs1020361	40,405,439	1.3	0.002	0.001	1:0	0.32
56	rs7725523	40,407,980	2.5	0.128	0.000	178:163	0.42
57	rs2120854	40,408,891	0.9	0.985	0.074	51:41	0.30
58	rs10473185	40,410,924	2.0	0.017	0.073	51:40	0.25
59	rs16869860	40,413,388	2.5	0.016	0.178	98:97	0.94
60	rs4957279	40,415,107	1.7	0.226	0.071	49:42	0.46
61	rs1031168	40,417,377	2.3	0.006	0.252	135:126	0.58
62	rs895123	40,419,818	2.4	0.000	0.442	143:116	0.093
63	rs1002922	40,422,312	2.5	0.020	0.314	155:144	0.52
64	rs1550761	40,424,886	2.6	0.009	0.202	89:88	0.94
65	rs2034182	40,426,689	1.8	0.015	0.074	51:39	0.21
66	rs4613763	40,428,485	1.8	0.087	0.159	98:97	0.94
67	rs1992662	40,429,609	1.1	1.000	0.000	155:143	0.49
68	rs10512736	40,431,143	1.5	0.171	0.313	155:143	0.49
69	rs16869934	40,433,109	2.0	0.078	0.254	110:101	0.54
70	rs10941505	40,435,003	1.9	0.022	0.456	192:176	0.40
71	rs1445009	40,437,446	2.4	0.002	0.105	53:47	0.55
72	***rs1025969***	***40,438,670***	***1.2***	***0.006***	***0.142***	***78:48***	***0.0075***
73	rs7723981	40,440,669	2.0	0.973	0.086	59:50	0.39
74	rs1445006	40,442,602	1.9	0.038	0.083	59:48	0.29
75	rs10512737	40,445,800	3.2	0.010	0.457	143:141	0.91
76	rs6451494	40,447,048	1.2	1.000	0.384	178:163	0.42
77	rs10473191	40,448,701	1.7	1.000	0.385	174:161	0.48
78	rs1992660	40,450,824	2.1	0.658	0.384	178:163	0.42
79	rs10071761	40,451,368	0.5	0.078	0.306	148:140	0.64
80	***rs4957288***	***40,459,597***	***8.2***	***0.045***	***0.153***	***83:56***	***0.022***
81	rs1899983	40,460,734	1.1	0.145	0.392	178:167	0.55
82	rs1445003	40,462,533	1.8	0.008	0.086	59:51	0.45
83	rs10512738	40,464,558	2.0	0.196	0.081	51:51	1.00
84	rs1545334	40,465,875	1.3	0.983	0.310	148:143	0.77
85	rs4957291	40,467,706	1.8	0.698	0.306	148:141	0.68
86	rs6878213	40,469,333	1.6	0.698	0.387	173:166	0.70
87	rs12514415	40,470,793	1.5	0.087	0.306	148:141	0.68
88	rs7734434	40,472,455	1.7	0.125	0.165	100:99	0.94
89	rs9292777	40,473,705	1.3	0.511	0.387	174:166	0.66
90	rs7722414	40,477,100	3.4	0.245	0.448	189:179	0.60
91	rs9292778	40,479,053	2.0	1.000	0.317	158:148	0.57
92	rs11957215	40,481,438	2.4	0.000	0.317	158:148	0.57
93	rs11954069	40,487,990	6.6	0.000	0.416	126:125	0.95
94	rs4498294	40,489,279	1.3	0.998	0.319	158:147	0.53
95	rs12520532	40,491,202	1.9	1.000	0.319	158:149	0.61
96	rs11955354	40,493,216	2.0	0.181	0.319	159:149	0.57
97	rs6451507	40,494,445	1.2	0.181	0.440	150:137	0.44
98	rs7725387	40,495,506	1.1	0.978	0.319	159:149	0.57
99	rs10941509	40,498,115	2.6	0.978	0.314	159:147	0.49
100	rs4957300	40,499,496	1.4	0.998	0.000	159:149	0.57
101	rs4957303	40,502,034	2.5	0.998	0.318	159:148	0.53
102	rs4532599	40,503,029	1.0	0.998	0.319	158:148	0.57
103	rs4495224	40,513,272	10.2	0.988	0.319	159:149	0.57
104	rs6871834	40,515,944	2.7	0.227	0.317	159:150	0.61

#	dbSNP ID	Position	Distance [kb]	LD [r²]	MAF	T:U	PTDT
105	rs6882351	40,517,411	1.5	0.226	0.095	59:56	0.78
106	rs10941515	40,518,283	0.9	0.000	0.311	155:148	0.69
107	rs7725052	40,523,027	4.7	0.001	0.000	122:109	0.39
108	rs10073277	40,524,443	1.4	0.749	0.310	160:139	0.22
109	***rs9283753***	***40,526,366***	***1.9***	***0.443***	***0.376***	***184:147***	***0.042***
110	***rs10041894***	***40,528,991***	***2.6***	***0.924***	***0.220***	***136:91***	***0.0028***
111	***rs12520522***	***40,530,435***	***1.4***	***0.766***	***0.209***	***135:90***	***0.0027***
112	***rs1505992***	***40,534,334***	***3.9***	***0.917***	***0.257***	***152:109***	***0.0078***
113	***rs1553575***	***40,538,689***	***4.4***	***1.000***	***0.000***	***154:112***	***0.010***
114	***rs1157160***	***40,540,177***	***1.5***	***0.984***	***0.263***	***149:109***	***0.013***
115	***rs4957311***	***40,542,168***	***2.0***	***0.984***	***0.267***	***155:113***	***0.010***
116	***rs7734013***	***40,543,000***	***0.8***	***1.000***	***0.263***	***154:112***	***0.010***
117	***rs1553577***	***40,545,764***	***2.8***	***1.000***	***0.263***	***154:112***	***0.010***
118	***rs7703539***	***40,549,071***	***3.3***	***1.000***	***0.264***	***154:111***	***0.0083***
119	***rs4957313***	***40,551,279***	***2.2***	***1.000***	***0.263***	***153:112***	***0.012***
120	***rs10068204***	***40,555,108***	***3.8***	***0.494***	***0.263***	***154:112***	***0.010***
121	***rs1876143***	***40,557,405***	***2.3***	***0.002***	***0.150***	***103:70***	***0.012***
122	rs1876142	40,557,462	0.1	0.013	0.294	126:114	0.44
123	***rs10941516***	***40,557,969***	***0.5***	***0.393***	***0.476***	***194:149***	***0.015***
124	***rs1876141***	***40,561,492***	***3.5***	***1.000***	***0.263***	***154:112***	***0.010***
125	***rs6873494***	***40,563,883***	***2.4***	***0.896***	***0.263***	***154:110***	***0.0068***
126	***rs4512170***	***40,566,291***	***2.4***	***0.551***	***0.243***	***142:103***	***0.013***
127	***rs4957319***	***40,568,372***	***2.1***	***0.551***	***0.150***	***103:72***	***0.019***
128	***rs6451517***	***40,571,930***	***3.6***	***0.000***	***0.243***	***142:103***	***0.013***
129	rs6451519	40,576,626	4.7	0.000	0.252	105:104	0.94
130	rs7730267	40,584,302	7.7	0.018	0.092	63:55	0.46
131	***rs969258***	***40,585,511***	***1.2***	***0.007***	***0.150***	***103:72***	***0.019***
132	rs7713972	40,588,231	2.7	0.002	0.000	114:110	0.79
133	rs6451525	40,590,026	1.8	0.009	0.000	32:19	0.069
134	***rs12697413***	***40,592,084***	***2.1***	***0.340***	***0.231***	***134:95***	***0.010***
135	rs1024098	40,595,144	3.1	0.002	0.092	63:55	0.46
136	rs6864749	40,604,350	9.2	0.002	0.000	95:87	0.55
137	***rs10055860***	***40,605,710***	***1.4***	***0.997***	***0.242***	***140:103***	***0.018***
138	***rs11742349***	***40,608,200***	***2.5***	***0.250***	***0.242***	***139:102***	***0.017***
139	rs2055646	40,610,699	2.5	0.031	0.074	52:38	0.14
140	rs6451529	40,611,106	0.4	0.005	0.000	65:58	0.53
141	***rs1948902***	***40,613,715***	***2.6***	***0.994***	***0.242***	***139:103***	***0.021***
142	***rs7708395***	***40,615,628***	***1.9***	***0.330***	***0.241***	***139:103***	***0.021***
143	rs4342367	40,619,398	3.8	0.988	0.095	68:53	0.17
144	rs7705019	40,623,561	4.2	0.752	0.095	68:53	0.17
145	rs906064	40,625,774	2.2	0.752	0.072	50:38	0.20
146	rs7711611	40,628,525	2.8	0.964	0.095	68:53	0.17
147	rs6878946	40,631,161	2.6	0.017	0.098	69:54	0.18
148	rs6869535	40,633,375	2.2	0.486	0.134	82:68	0.25
149	***rs6897467***	***40,635,001***	***1.6***	***0.963***	***0.240***	***137:104***	***0.034***
150	***rs906062***	***40,636,914***	***1.9***	***0.200***	***0.247***	***140:106***	***0.030***
151	rs4409138	40,639,362	2.4	0.204	0.000	57:50	0.50
152	***rs6879632***	***40,640,517***	***1.2***	***0.375***	***0.241***	***137:105***	***0.040***
153	rs6890899	40,642,798	2.3	0.005	0.107	76:58	0.12
154	rs16870171	40,643,917	1.1	0.005	0.043	29:21	0.26
155	rs4957324	40,647,732	3.8	0.014	0.107	76:58	0.12
156	rs924967	40,650,879	3.1	0.060	0.000	128:121	0.66
157	***rs1979937***	***40,653,159***	***2.3***	***0.378***	***0.247***	***141:109***	***0.043***
158	***rs10941519***	***40,655,512***	***2.4***	***0.140***	***0.464***	***191:143***	***0.0086***
159	rs4957137	40,656,550	1.0	0.004	0.108	78:58	0.086
160	rs17834885	40,661,879	5.3	0.004	0.192	107:100	0.63
161	rs17238368	40,662,946	1.1	0.009	0.108	62:45	0.10
162	rs11739531	40,665,104	2.2	0.135	0.332	150:148	0.91
163	rs6892418	40,667,478	2.4	0.223	0.214	125:115	0.52
164	rs1395082	40,668,032	0.6	0.005	0.452	175:170	0.79
165	rs7714305	40,677,005	9.0	0.000	0.213	95:83	0.37
166	rs7716982	40,682,317	5.3	0.409	0.331	149:148	0.95
167	rs10941521	40,683,897	1.6	1.000	0.452	175:170	0.79
168	rs10068069	40,687,372	3.5	1.000	0.452	175:170	0.79
169	rs10071870	40,688,984	1.6	0.998	0.452	175:170	0.79
170	rs4957334	40,692,609	3.6	0.998	0.452	171:167	0.83
171	rs11744569	40,693,406	0.8	0.409	0.452	175:170	0.79
172	rs2888173	40,695,209	1.8	0.001	0.331	149:149	1.00
173	rs7717878	40,698,014	2.8	0.000	0.001	1:0	0.32
174	rs1395085	40,701,679	3.7	0.147	0.153	95:89	0.66
175	rs10473214	40,703,792	2.1	0.007	0.452	177:169	0.67
176	rs1553579	40,704,550	0.8	0.007	0.212	92:80	0.36
177	rs1553578	40,704,941	0.4	0.006	0.454	176:170	0.75
178	rs4432939	40,706,856	1.9	0.003	0.315	116:108	0.59
179	rs4607392	40,709,272	2.4	0.476	0.367	164:161	0.87
180	rs10066459	40,711,291	2.0	0.472	0.450	179:171	0.67
181	rs4434423	40,713,439	2.1	0.995	0.366	164:160	0.82
182	rs4133101	40,715,324	1.9	0.039	0.367	164:159	0.78
183	rs2228058	40,717,011	1.7	0.043	0.063	41:39	0.82
184	rs4546432	40,718,307	1.3	0.678	0.000	173:166	0.70
185	rs11957406	40,719,324	1.0	0.478	0.485	178:170	0.67
186	rs6451535	40,723,788	4.5	0.137	0.000	154:145	0.60
187	rs17835256	40,726,418	2.6	0.046	0.233	132:127	0.76
188	rs16870224	40,728,697	2.3	0.073	0.140	92:83	0.50
189	rs4277957	40,735,275	6.6	0.767	0.309	156:147	0.61
190	rs1057631	40,748,001	12.7	1.000	0.256	134:127	0.66
191	rs10737963	40,751,910	3.9	0.017	0.256	134:127	0.66
192	rs249413	40,757,195	5.3	0.031	0.162	79:77	0.87
193	rs447810	40,762,558	5.4	0.138	0.312	156:149	0.69
194	rs3805499	40,764,958	2.4	0.138	0.233	132:127	0.76
195	rs249414	40,769,451	4.5	0.335	0.312	156:149	0.69
196	rs10941525	40,771,683	2.2	0.991	0.424	181:168	0.49
197	rs7721093	40,774,697	3.0	0.335	0.426	181:168	0.49
198	rs171605	40,776,880	2.2	1.000	0.310	155:146	0.60
199	rs249422	40,778,299	1.4	0.334	0.310	155:146	0.60
200	rs7719692	40,779,729	1.4	1.000	0.426	182:169	0.49
201	rs3792827	40,781,090	1.4	0.260	0.426	182:169	0.49
202	rs3805497	40,782,642	1.6	0.158	0.261	138:130	0.63
203	rs2329353	40,784,025	1.4	0.985	0.319	130:125	0.75
204	rs249425	40,788,522	4.5	0.010	0.310	155:147	0.65
205	rs29743	40,791,037	2.5	0.004	0.353	138:118	0.21
206	rs257009	40,806,973	15.9	0.000	0.160	76:64	0.31
207	rs3805492	40,819,020	12.0		0.381	142:121	0.20