



- (51) **International Patent Classification:**  
*C12Q 1/68* (2006.01) *G01N 33/48* (2006.01)
- (21) **International Application Number:**  
PCT/AU2012/000440
- (22) **International Filing Date:**  
27 April 2012 (27.04.2012)
- (25) **Filing Language:** English
- (26) **Publication Language:** English
- (30) **Priority Data:**  
61/480,987 29 April 2011 (29.04.2011) US
- (71) **Applicants (for all designated States except US):** **THE UNIVERSITY OF SYDNEY** [AU/AU]; Sydney, New South Wales 2006 (AU). **WESTERN SYDNEY LOCAL HEALTH DISTRICT** [AU/AU]; CNR Hawkesbury & Darcy Road, Westmead, Sydney, New South Wales 2145 (AU). **THE WALTER AND ELIZA HALL INSTITUTE OF MEDICAL RESEARCH** [AU/AU]; 1G Royal Parade, Parkville, Melbourne, Victoria 3052 (AU). **MURDOCH UNIVERSITY** [AU/AU]; South Street, Murdoch, Perth, Western Australia 6150 (AU).
- (72) **Inventors; and**
- (75) **Inventors/Applicants (for US only):** **GEORGE, Jacob** [AU/AU]; 9 Sunridge Place, West Pennant Hills, New South Wales 2125 (AU). **BOOTH, David** [AU/AU]; 8 Beresford Avenue, Chatswood, New South Wales 2067 (AU). **SUPPIAH, Vijayaprakash** [SG/AU]; 40a De-spointes Street, Marrickville, New South Wales 2204 (AU). **STEWART, Graeme** [AU/AU]; 52 Pacific Road, Palm Beach, New South Wales 2108 (AU). **GAUDIERI, Silvana** [AU/AU]; 10 Sawle Road, Hamilton Hill, Western Australia 6163 (AU). **MALLAL, Simon** [AU/AU]; 93

Evandale Street, Floreat, Western Australia 6014 (AU). **SMITH, Katherine** [AU/AU]; c/o- 1G Royal Parade, Parkville, Victoria 3052 (AU). **BAHLO, Melanie** [AU/AU]; c/o- 1G Royal Parade, Parkville, Victoria 3052 (AU).

- (74) **Agents:** **OLIVE, Mark** et al.; c/- FB Rice, Level 23, 44 Market Street, Sydney, New South Wales 2000 (AU).
- (81) **Designated States (unless otherwise indicated, for every kind of national protection available):** AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, 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, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, 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, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

**Published:**

- with international search report (Art. 21(3))
- with sequence listing part of description (Rule 5.2(a))

(54) **Title:** METHOD OF DETERMINING RESPONSE TO TREATMENT WITH IMMUNOMODULATORY COMPOSITION

(57) **Abstract:** The present invention provides a method for accurately determining the likelihood that a subject will respond to treatment with an immunomodulatory composition comprising detecting homozygosity for HLA-C2 alleles or at least two markers one of which is a marker for an HLA-C allele and one of which is a marker for an IL28B allele, and processes for selecting suitable subjects for therapy or for continued therapy, and for providing appropriate therapy to subjects, based on the assay results.



Method of determining response to treatment with immunomodulatory composition

Cross reference to related application

This application claims priority to United States application No. 61/480,987 filed April  
5 29, 2011 the full contents of which is incorporated by reference herein.

Field of the invention

The present invention is in the field of diagnostic and prognostic assays for medical  
conditions that are treated using an immunomodulatory composition, and improved  
10 therapeutic methods based on the diagnostic and prognostic assays of the invention.

Background to the invention

*Mechanism of action of immunomodulatory compositions*

Immunomodulatory compositions comprise drug compounds that act by modulating  
15 certain key aspects of the immune system in the treatment of viral diseases, neoplasias,  
Th1-mediated diseases, Th2-mediated diseases, or Th17-mediated diseases,  
substantially by modulating expression or secretion of one or more cytokines involved  
in autoimmunity and/or immune responses to infectious agents, or by modulating one  
of more components of a cytokine signalling pathway.

20

Cytokines may be interferons (IFNs, *e.g.*, Type I IFNs such as IFN- $\alpha$ , IFN- $\beta$ , or IFN- $\omega$ ;  
or Type II IFNs such as IFN- $\gamma$ ; or Type III IFNs such as IFN- $\lambda$ 1, IFN- $\lambda$ 2, or IFN- $\lambda$ 3),  
interleukins (*e.g.*, IL-1, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-10, IL-11, IL-  
12, IL-13, IL-14, IL-15, IL-16, IL-17, IL-18, IL-21, or IL-35), a tumor necrosis factor  
25 (*e.g.*, TNF- $\alpha$  or TNF- $\beta$ ), or colony-stimulating factor (CSF). The IFNs generally assist  
immune responses by inhibiting viral replication within host cells, activating cytotoxic  
T cells and macrophages, increasing antigen presentation to lymphocytes, inducing  
resistance to viral and intracellular bacterial infections, and controlling tumors.  
Additionally, the Type III IFNs exert a regulatory effect on Th2 cells. Interleukins  
30 promote development and differentiation of T cells, B cells and hematopoietic cells.  
Tumor necrosis factors regulate cells of the immune systems to stimulate acute phase

inflammatory responses, induce apoptotic cell death, inhibit tumorigenesis and inhibit viral replication. Although IFNs may be produced by a number of different cells, IFN- $\gamma$  is produced predominantly by Th1 cells, and interleukins and TNF- $\alpha$  are produced by Th1 cells and/or Th2 cells.

5

Th1 cells and Th2 cells are effector T cells defined by their cytokine secretion profiles. Th1 cells mediate cellular immunity to protect against intracellular pathogens and immunogens via the actions of cytotoxic T lymphocytes and activated macrophages and complement-fixing and complement-opsonizing antibodies. Th1 cells produce IL-  
10 2, which stimulates growth and differentiation of T cell responses mediated by Th1 cells, as well as producing IFN- $\gamma$  and TNF- $\beta$ . On the other hand, Th2 cells mediate humoral immunity and allergic responses to protect against extracellular pathogens and antigens via the actions of B cells, mast cells and eosinophils. Th2 cells produce IL-3, IL-4, IL-5, and IL-10, which stimulate production of IgE antibodies, and also  
15 recruitment, proliferation, differentiation, maintenance and survival of eosinophils.

Certain Th1-mediated and Th2-mediated diseases are driven by disruption of the balance between Th1 cells and Th2 cells. The finely-tuned balance of Th1 and Th2 cells is regulated by cytokine secretion and, under normal circumstances, Th2 cells  
20 secrete IL-4 and IL-10 which down-regulate Th1 cells thereby regulating production of IFN- $\gamma$ , TNF- $\beta$  and IL-2. In particular, IL-10 is a potent inhibitor of Th1 cells. IFNs such as IFN- $\gamma$  also drive Th1 cell production. Conversely, IL-4 drives Th2 cell production and IFN- $\gamma$  inhibits Th2 cells. In Th1-mediated diseases *e.g.*, multiple sclerosis (MS), rheumatoid arthritis (RA), Type I diabetes (IDDM) and scleroderma,  
25 delayed type hypersensitivity (DTH) occurs in those organ systems in which CD4<sup>+</sup> Th1 cells are over-activated relative to Th2 cells. In MS, a Th1/Th2 imbalance in the central nervous system leads to proliferation of pro-inflammatory CD4<sup>+</sup> Th1 cells, IFN- $\gamma$  secretion, macrophage activation and consequential immune-mediated injury to myelin and oligodendrocytes, wherein the IFN- $\gamma$  release in this case may also drive Th1  
30 cell overproduction. In IDDM, a Th1/Th2 imbalance occurs in the thymus and periphery leading to progressive elimination of functional Th2 cells as autoreactive Th1

cells become activated and mediate pancreatic islet  $\beta$ -cell destruction. In localized scleroderma, the administration of IL-12 may restore Th1/Th2 immune balance. In contrast, Th2-mediated diseases *e.g.*, Con A hepatitis, atopic dermatitis, asthma and allergy, are generally characterized by over-production of IgE antibodies and/or eosinophilia as a consequence of a Th1/Th2 imbalance. In Con A hepatitis, repeated injections of Con A shift an initial Th1 response to a Th2 and profibrogenic response, with over-production and secretion of IL-4, IL-10 and TGF- $\beta$  in the liver activating natural killer T cells as part of an innate immune response thereby causing liver damage.

10

Th17 cells provide an effector arm distinct from Th1 and Th2 cells and, like Treg (iTreg), are regulated by TGF- $\beta$ . Th17 cellular differentiation is important for host defense *e.g.*, against bacteria and fungi, and poor regulation of Th17 cellular function is implicated in immune pathogenesis of autoimmune and inflammatory diseases.

15

*Medical indications for immunomodulatory compositions*

Infections by a number of different viruses are treated using immunomodulatory compositions, including infections by human papillomaviruses such as HPV16, HPV6, HPV11; infections by herpesviruses such as HSV-1, HSV-2, VZV, HHV-6, HHV-7, HHV-8 (KSHV), HCMV and EBV; infections by picornaviruses such as the Coxsackie B viruses and encephalomyocarditis virus (EMCV); infections by flaviviruses such as the encephalitis viruses and hepatitis viruses *e.g.*, hepatitis A virus, hepatitis B virus (HBV) and hepatitis C virus (HCV); arenaviruses such as those associated with a viral haemorrhagic fever; infections by togaviruses such as equine encephalitis viruses; infections by bunyaviruses such as Rift Valley fever virus, Crimean-Congo haemorrhagic fever virus, Hantaan hantavirus (HTNV) and Apeu virus (APEUV); infections by filoviruses such as Ebola virus and Marburg virus; infections by paramyxoviruses such as respiratory syncytial virus (RSV); infections by rhabdoviruses such as vesicular stomatitis virus (VSV); infections by orthomyxoviruses such as the influenza viruses *e.g.*, influenza A virus (IAV); and infections by coronaviruses such as SARS-associated coronavirus. Neoplasias are also treated using immunomodulatory

30

compositions *e.g.*, HPV-associated cancer such as cervical intraepithelial neoplasia, cervical carcinoma, vulvar intraepithelial neoplasia, penile intraepithelial neoplasia, perianal intraepithelial neoplasia; hepatocellular carcinoma; basal cell carcinoma, squamous cell carcinoma, actinic keratosis, and melanoma. Certain Th2-mediated  
5 diseases *e.g.*, asthma, allergic rhinitis, atopic dermatitis, are also treated using immunomodulatory compositions.

*Cytokines as immunomodulatory compositions*

Partly by virtue of the modulation of cytokines and cytokine signalling by  
10 immunomodulatory compositions, it is known to use cytokines *per se* as immunomodulatory compositions.

For example, IFNs in general possess antiviral and anti-oncogenic properties, the ability to stimulate macrophage and natural killer cell activation, and the ability to  
15 enhance MHC class I and II molecules for presentation of foreign peptides to T cells. In many cases, the production of IFNs is induced in response to infectious agents, foreign antigens, mitogens and other cytokines *e.g.*, IL-1, IL-2, IL-12, TNF and CSF. Thus, IFNs and IFN inducers have gained acceptance as therapeutic agents in the treatment of infections, neoplasias, Th1-mediated disease and Th2-mediated disease.  
20 IFNs are known to be used for treatment of infections by several positive-sense single-stranded RNA viruses *i.e.*, (+) ssRNA viruses, including *e.g.*, SARS-associated coronavirus, HBV, HCV, coxsackie B virus, EMCV, and for treatment of infections by several negative-sense single-stranded RNA viruses *i.e.*, (-) ssRNA viruses, including  
*e.g.*, Ebola virus, VSV, IAV, HTNV and APEUV (*see e.g.*, De Clercq *Nature Reviews* 2,  
25 704-720 (2004); Li *et al.*, *J. Leukocyte Biol.*, online publication DOI:10.1189/jlb.1208761 (Apr 30, 2009). In such formulations, the IFN, especially IFN- $\alpha$  may be pegylated. Pegylated IFN- $\lambda$ 1 is currently in clinical trial for treatment of chronic HCV infection, and has been shown to be useful for protecting isolated cells against VSV, EMCV, HTNV, APEUV, IAV, HSV-1, HSV-2 and HBV, (*see e.g.*, Li *et al.*,  
30 *J. Leukocyte Biol.*, online publication DOI:10.1189/jlb.1208761, Apr 30, 2009). IFN- $\alpha$  is also known to be used in the treatment of certain lesions and neoplasias *e.g.*,

condylomata acuminata, hairy cell leukemia, Kaposi's sarcoma, melanoma, non-Hodgkin's lymphoma, however IFN- $\beta$  has been shown to have potent anti-tumor activity against human astrocytoma/glioblastoma cells, whereas IFN- $\lambda$ 1 has been shown to have activity against glioblastoma cells, thymoma cells and fibrosarcoma cells, and IFN- $\lambda$ 2 has been shown to have activity against melanoma and fibrosarcoma cells (see e.g., Li et al, *J. Leukocyte Biol.*, online publication DOI:10.1189/jlb.1208761, Apr 30, 2009). It is also known to use IFN- $\beta$  for treatment of relapsing forms of Th1-mediated diseases such as MS. IFN- $\lambda$ 2 has also been shown to protect against certain Th2-mediated diseases e.g., asthma and Con A-induced hepatitis (see e.g., Li et al, *J. Leukocyte Biol.*, online publication DOI:10.1189/jlb.1208761, Apr 30, 2009).

Since the expression of all IFN- $\lambda$  proteins are induced by IFN- $\alpha$ , IFN- $\beta$  and IFN- $\lambda$  molecules e.g., Sirén et al., *J. Immunol.* 174, 1932-1937 (2005), Ank et al., *J. Virol* 80, 4501-4509 (2006) and Ank et al., *J. Immunol.* 180, 2474-2485 (2008), immunomodulatory compositions comprising IFN- $\alpha/\beta$  may act, at least in part, to induce IFN- $\lambda$  proteins as effector molecules. The receptor complex for Type I IFNs consists of a heterodimeric IFNAR1/IFNAR2 complex, whereas Type III IFNs signal through a heterodimeric IL-28R $\alpha$ /IL-10R2 receptor e.g., Li et al, *J. Leukocyte Biol.*, online publication DOI:10.1189/jlb.1208761 (Apr 30, 2009). IL-28R $\alpha$ /IL-10R2 is expressed in far fewer contexts than IFNAR1/IFNAR2. This suggests that therapy using immunomodulatory compositions comprising IFN- $\alpha/\beta$  may be less specific than therapy using immunomodulatory compositions comprising IFN- $\lambda$ . For example, administration of IFN- $\alpha/\beta$  may activate both receptor types i.e., directly via action of IFN- $\alpha/\beta$  on IFNAR1/IFNAR2 receptors and indirectly via induction of IFN- $\lambda$  and subsequent action of IFN- $\lambda$  on IL-28R $\alpha$ /IL-10R2 receptors. Conversely, administration of IFN- $\lambda$  is likely to activate selectively IL-28R $\alpha$ /IL-10R2 receptors. Notwithstanding that this may be the case, all IFNs activate the Jak/STAT pathways and generally induce common interferon-stimulated genes (ISGs) that mediate the biological effects of IFNs e.g., Sirén et al., *J. Immunol.* 174, 1932-1937 (2005), Ank et al., *J. Virol* 80, 4501-4509 (2006), Ank et al., *J. Immunol.* 180, 2474-2485 (2008), and Li et al, *J. Leukocyte Biol.*, online publication DOI:10.1189/jlb.1208761 (Apr 30, 2009).

*Cytokine-inducing agents as immunomodulatory compositions*

Various immunomodulatory compositions that induce IFN production *e.g.*, poly(I)-poly(C), poly(I)-poly(C<sub>12</sub>-U) or ampligen, and deazaneplanocin A, are also used in the treatment of infections by *e.g.*, coxsackie B virus, Ebola virus and for certain flaviviruses and bunyaviruses that are amenable to treatment with IFNs (De Clerq *Nature Reviews* 2, 704-720 (2004). Immunomodulatory compounds may also exert their activity by activating Toll-like receptors (TLRs) to induce selected cytokine biosynthesis.

10

*Immunomodulatory nucleotide analogs*

Immunomodulatory guanosine analogs, such as those having substituents at the 7-position and/or 8-position, *e.g.*, Reitz *et al.*, *J. Med. Chem.* 37, 3561-3578 (1994) Michael *et al.*, *J. Med. Chem.* 36, 3431-3436 (1993) have been shown to stimulate the immune system, whilst 5'-O-propionyl and 5'-O-butyryl esters of 2-amino-6-methoxy-9-( $\beta$ -D-arabinofuranosyl)-9H-purine inhibit varicella zoster virus (VZV) *e.g.*, U.S. Pat. No. 5,539,098 to Krenitsky. Other guanosine analogs, in particular 6-alkoxy derivatives of arabinofuranosyl purine, are useful for anti-tumor therapy *e.g.*, U.S. Pat. No. 5,821,236 to Krenitsky. The 7-deazaguanosine analogs have been shown to exhibit antiviral activity in mice against a variety of RNA viruses, whereas 3-deazaguanine analogs have significant broad spectrum antiviral activity against certain DNA and RNA viruses *e.g.*, Revankar *et al.*, *J. Med. Chem.* 27, 1489-1496 (1984), and certain 7-deazaguanine and 9-deazaguanine analogs protect against a lethal challenge of Semliki Forest virus *e.g.*, Girgis *et al.*, *J. Med. Chem.* 33, 2750-2755 (1990). Selected 6-sulfenamide and 6-sulfinamide purine nucleosides are also disclosed in U.S. Pat. No. 4,328,336 to Robins as having demonstrated significant antitumor activity. Wang *et al.* (WO 98/16184) also disclose purine L-nucleoside compounds and analogs thereof were used to treat an infection, infestation, a neoplasm, an autoimmune disease, or to modulate aspects of the immune system. Guanosine analogs *e.g.*, ribavirin and derivatives thereof *e.g.*, acetate salts or ribavirin 5'-monophosphate or ribavirin 5'-diphosphate or ribavirin 5'-triphosphate or ribavirin 3',5'-cyclic phosphate or the 3-

30

carboxamide derivative taribavirin (viramidine), 7-benzyl-8-bromoguanine, 9-benzyl-8-bromoguanine, and CpG-containing oligonucleotides, that shift the Th1/Th2 balance and are useful for the treatment of Th1-mediate or Th2-mediated disease depending upon their cytokine profiles. These compounds have been shown to elicit various effects on lymphokines IL-1, IL-6, IFN- $\alpha$  and TNF- $\alpha$  *e.g.*, Goodman, *Int. J. Immunopharmacol*, 10, 579-588 (1988); U.S. Pat. No. 4,746,651; Smee *et al.*, *Antiviral Res.* 15, 229 (1991); Smee *et al.*, *Antimicrobial Agents and Chemotherapy* 33, 1487-1492 (1989). For example, 7-benzyl-8-bromoguanine and 9-benzyl-8-bromoguanine selectively inhibit Th1 cytokine production, specifically IL-2 and IFN- $\gamma$  and therefore may be useful in the treatment of Th1-related autoimmune disease, which manifests activated T cells and overproduction of IFN- $\gamma$ , and target leukemia and lymphoma cells, *e.g.*, Poluektova *et al.*, *Int. J. Immunopharmacol.* 21, 777-792 (1999). In contrast, ribavirin shifts an immune response from Th2 toward a Th1 cytokine profile, and is useful for treatment of Th2-mediated diseases. Ribavirin is useful in post-exposure prophylaxis of exposure to *e.g.*, arenaviruses causing Lassa fever or Crimean-Congo hemorrhagic fever, HTNV, West Nile Virus, chronic HCV infection, AIV and RSV.

Various other immunomodulatory nucleotide analogs possess potent antiviral activity, and may restore p53 function in HPV-associated cancers *e.g.*, cidofovir [(S)-1-(3-hydroxy-2-phosphonylmethoxypropyl)cytosine, (HPMPC)] *e.g.*, Abdulkarin *et al.*, *Oncogene* 21, 2334-2346, (2002). Cidofovir is used in the treatment of a number of viral conditions including HCMV-retinitis in AIDS patients and other HCMV infections and poxvirus infections.

#### 25 *Other classes of immunomodulatory compositions*

Other classes of immunomodulatory compositions include small organic molecule imidazoquinoline amine derivatives *e.g.*, U.S. Pat. Nos. 4,689,338 and 6,069,149; purine derivatives *e.g.*, U.S. Pat. Nos. 6,028,076 and 6,376,501; imidazopyridine derivatives; *e.g.*, U.S. Pat. No. 6,518,265; benzimidazole derivatives *e.g.*, U.S. Pat. No. 6,387,938; adenine derivatives *e.g.*, U.S. Pat. No. 6,376,501; and 3- $\beta$ -D-ribofuranosylthiaz-olo[4,5-d]pyrimidine derivatives *e.g.*, U.S. Pat. publication No.

200301994618. The immunosuppressive agent mycophenolate mofetil inhibits coxsackie B3 virus-induced myocarditis (see, *e.g.*, Padalko *et al.*, *BMC Microbiol.* 3, 25 *et seq.* (2003)).

- 5 The list of immunomodulatory compositions provided herein is not exhaustive and a number of other compound classes are also known in the art *e.g.*, in U.S. Pat. Nos. 5,446,153; 6,194,425; and 6,110,929.

### *Contraindications*

- 10 Many immunomodulatory compositions produce adverse side-effects, suggesting a benefit in limiting their application to contexts where therapeutic benefit outweighs detrimental effects. In addition to the favourable changes in the immune system that immunomodulatory compositions produce in therapy, imbalances occur. For example, the IFNs may cause, *inter alia*, psychiatric disorders, depression, anaphylaxis, 15 thrombocytopenia, seizure, cardiomyopathy, hepatotoxicity, flu-like symptoms, fever, fatigue, headache, muscle pain, convulsions, dizziness, erythema and immunosuppression through neutropenia, and interleukins *e.g.*, IL-1, may cause dose-related fever and flu-like symptoms. In another example, guanosine analogs may be teratogenic with prolonged use.

20

- The efficacy of immunomodulatory compositions for particular indications may be highly variable, and therapeutic outcome is likely to be influenced by host factors *e.g.*, genotype. For example, racial differences may affect suitability of subjects for therapy with immunomodulators. In another example, HLA haplotype effects, governing both 25 innate and adaptive immune responses of subjects, are also known to affect viral clearance. However, because MHC class I proteins are a diverse class of heterodimeric receptors that present antigen fragments to cytotoxic T-cells via the CD8 receptor on the cytotoxic T-cells, and also bind to a diverse class of inhibitory and activating receptors on natural killer (NK) cells, interaction *e.g.*, an additive interaction or 30 epistatic interaction between expressed HLA Class I alleles can produce variations in

immune responses. HLA gene dosage effects can also be significant in determining innate immune responses of subjects.

Accordingly, means for identifying and selecting those patients who are likely to respond to treatment with an immunomodulatory composition may provide a substantial therapeutic benefit to those patients that are either non-responders, low-responders or relapsers, by avoiding inappropriate prescriptions to those patient classes and reducing the anxiety caused by subsequent treatment failure. More accurate prescription of drugs to responders also provides for reduced subsidies by health agencies. Moreover, for those conditions in which alternative therapies are available, such means may also provide for selection of the most appropriate therapy for a particular patient.

*Prognostic indicators for effective therapy with immunomodulatory compositions*

15 KIR2/HLA-C genotypes

Khakoo *et al.*, *Science* 305, 872 – 874, 2004 have demonstrated that genes encoding the natural killer (NK) cell immunoglobulin-like receptor (KIR) isoform KIR2DL3 and its human leukocyte antigen C group 1 (HLA-C1) ligand influence spontaneous resolution of hepatitis C virus (HCV) infection in Caucasians and African Americans having low infectious doses of HCV, but not in subjects having a high-dose exposure. Khakoo *et al.*, suggested that, for subjects having a high-dose exposure, the innate immune response mediated by NK cells is likely overwhelmed. Thus, homozygosity for KIR2DL3 and the cognate HLA-C1 ligand is known to be associated with enhanced probability of spontaneous HCV clearance via generation of an innate immunity to the virus.

The HLA-C gene locus encodes ligands for the natural killer (NK) cell immunoglobulin-like receptor-2 isoforms (KIR2DL), wherein a functional dimorphism determines KIR2DL specificity such that HLA-C group 1 (HLA-C1) alleles encoding HLA-C comprising Asn80 bind to the inhibitory receptors KIR2DL2 and KIR2DL3 and to the activating KIR2DS2 receptor, and HLA-C group 2 (HLA-C2) alleles

encoding HLA-C comprising Lys80 bind to the inhibitory receptor KIR2DL1 and to the activating receptor KIR2DS1.

Sequences of HLA-C molecules that permit their classification into HLA-C1 or HLA-C2 genotypes are provided by the IMGT/HLA database of the European Bioinformatics Institute (EBI), Wellcome Trust, Genome Campus, Hinxton, Cambridge CB10 1SD, United Kingdom. *See e.g., Robinson et al., Nucleic Acids Res. 31, 311-314 (2003).* Based on primary amino acid sequence data available in the IMGT/HLA database as at 7 February, 2011, exemplary HLA-C1 molecules include most HLA-Cw1, HLA-Cw3, HLA-Cw7, HLA-Cw8, HLA-Cw12, HLA-Cw13, HLA-Cw14 and HLA-Cw16 genotypes, in addition to HLA-Cw\*0227 and HLA-Cw\*0411, HLA-Cw\*0429, HLA-Cw\*0436, HLA-Cw\*0455, HLA-Cw\*0611, HLA-Cw\*1507, HLA-Cw\*1525, and HLA-Cw\*1543. Also based on those sequence data, exemplary HLA-C2 molecules include most HLA-Cw2, HLA-Cw4, HLA-Cw5, HLA-Cw6, HLA-Cw15, HLA-Cw17 and HLA-Cw18 genotypes, in addition to HLA-Cw\*0114, HLA-Cw\*0307, HLA-Cw\*0315, HLA-Cw\*0345, HLA-Cw\*0707, HLA-Cw\*0709, HLA-Cw\*0776, HLA-Cw\*0810, HLA-Cw\*1204, HLA-Cw\*1205, HLA-Cw\*1209, HLA-Cw\*1221, HLA-Cw\*1233, HLA-Cw\*1241, HLA-Cw\*1412, HLA-Cw\*1602, HLA-Cw\*1609, HLA-Cw\*1612, HLA-Cw\*1619, and HLA-Cw\*1625.

20

Romero *et al.*, *Mol Immunol.*, 45, 2429-2436, 2008 confirmed the KIR2/HLA-C interaction *e.g.*, an additive or epistatic interaction between KIR2DL3 homozygosity and homozygosity in HLA-C1 in enhancing the likelihood of spontaneous clearance of HCV in a cohort of 160 Puerto-Rican American drug users with HCV. Romero *et al.* also found a new interaction *e.g.*, an additive interaction or epistatic interaction between KIR2DL3 and HLA-DRB1\*1201 is associated with spontaneous clearance

Knapp *et al.*, *Heptol.* 51, 1168-1175, 2009 sought to determine whether or not the KIR2DL3/HLA-C1 gene combination provides an advantage in HCV-exposed injection drug user having apparent resistance to HCV infection *i.e.*, because they remain seronegative and aviremic, and in chronically-infected subjects who

successfully clear HCV with treatment. Knapp *et al.* found a stronger association between homozygosity for KIR2DL3 in combination with a HLA-C1 allotype in exposed seronegative aviremic subjects than for subjects with chronic HCV infection. Knapp *et al.* also concluded that homozygosity for KIR2DL3 in combination with a  
5 HLA-C1 allotype was indicative of an improved response to antiviral therapy as determined by sustained virological response (SVR).

In a more recent study, Vidal-Castineira *et al.*, *J Virol.* 84, 475-481, 2010 examined the effects of different combinations of KIRs with their HLA class I ligands on the  
10 response to combined treatment for HCV infection using pegylated IFN $\alpha$  and ribavirin, and concluded that the frequency of the KIR2DL2 allele was significantly increased in non-responder (NR) subjects, whereas the KIR2DL3 allele was significantly increased in the SVR, and that the frequency of homozygosity for the KIR2DL2 allele when combined with the specific heterozygous genotype HLA-C1/HLA-C2 was significantly  
15 increased in NR subjects, whereas homozygosity for both KIR2DL3 and HLA-C1 alleles is more closely associated with SVR.

The highly polymorphic nature of interactions *e.g.*, an additive interaction or epistatic interaction between HLA and KIR2DL receptors, and our limited understanding of the  
20 mechanisms by which genetic variations at the HLA and KIR2DL loci affect therapeutic responses, suggest that the relevance of KIR2/HLA-C genotype combinations does not indicate an association for HLA-C alleles in isolation. For example, effects at the HLA-C1 or HLA-C2 loci are not discernable unless evaluated in combination with specific KIR2DL genotypes, because several disclosures suggest  
25 non-significance in the association for HLA-C *per se* with response to treatment with ribavirin and pegylated INF- $\alpha$ . See, *e.g.*, Carniero *et al.*, *Liver International*, 567-572 (2010), at page 570, left col., lines 2-10, and at page 572, left col., lines 25-28 and Figure 2; and Vidal-Castineira *et al.*, *J. Virol* 84, 475-481 (2010), at Table 2 on page 478. Singh *et al.*, *World J. Gastroenterol.* 13, 1770-1787 (2007) suggested that the HLA  
30 Class II molecule HLA-DR is a prominent immunogenetic factor influencing interferon treatment response.

### IL28B genotypes and haplotypes

International Patent Application No. PCT/AU2010/000713 (WO 2010/144946), incorporated herein in its entirety by way of reference, describes several SNP markers  
5 linked to the IL28B gene for determining the likelihood of a response of an HCV-infected subject to therapy with an immunomodulatory composition, including the SNPs rs8099917, rs4803221, and rs1297860. The marker rs8099917 is accepted in the art as a highly-predictive marker of the response of HCV-infected subjects to standard therapy with ribavirin and pegylated IFN- $\alpha$ . Notwithstanding the availability of  
10 rs8099917, the availability of reliable tests remains limited, such as tests that have high predictive value and/or specificity and/or provide population coverage of more than about 10%.

### Summary of the invention

#### 15 1. *Introduction*

In work leading to the present invention, the inventors sought to develop improved tests for determining a response of individuals with chronic HCV infection to treatment with immunomodulatory compositions such as comprising IFNs, specifically pegylated IFN- $\alpha$ , in combination with ribavirin. The inventors sought to ascertain new genotypes,  
20 serotypes, haplotypes and genotypes within the IL28B gene, and epistatic interactions between such IL28B alleles and alleles at other loci that are associated with viral clearance in response to treatment.

More particularly, the inventors sought to ascertain additive and/or epistatic  
25 interactions between HLA Class I alleles and IL28B alleles that provide a basis for improved tests for determining a response *e.g.*, a low response (LR) or high response (HR), to antiviral therapy. In one approach, the inventors sought to determine whether or not a combination of one or more HLA-C2 alleles and one or more IL28B alleles in a subject's genome is predictive of a response *e.g.*, a low response (LR) or high  
30 response (HR) to antiviral therapy. In another example, the present inventors sought to determine whether or not a combination of one or more HLA-C1 alleles and one or

more IL28B alleles is predictive of a response *e.g.*, a low response (LR) or high response (HR) to antiviral therapy.

The inventors performed HLA-C genotyping and serotyping and IL28B genotyping on  
5 a cohort of about 910 HCV-infected subjects, and determined associations between HLA-C alleles and/or IL28B alleles and responsiveness to therapy. The inventors were particularly interested in determining single alleles, and combinations of IL28B alleles with HLA-C alleles, that provide an improved predictor of the response to therapy relative to known tests.

10

In summary, the inventors found that HLA-C2 homozygosity when combined with one or more copies of a low response (LR) allele of the SNP rs7248668, an allele linked to IL28B, *e.g.*, the A allele of rs7248668, in a subject's genome is highly predictive of a low response or poor response or non-response to antiviral therapy with  
15 immunomodulatory agent(s). The predictive capacity of the specific HLA-C2/IL28B-LR allele combination is especially evident from a comparison of the treatment responses of HLA-C2 homozygotes carrying one rs7248668 LR allele in his/her genome to the treatment responses of subjects carrying at least one HLA-C1 allele and at least one rs7248668 LR allele and/or to the treatment responses of subjects two  
20 copies of the corresponding high response (HR) allele of the SNP *e.g.*, the G allele of rs7248668. For example, data presented herein demonstrate a positive predictive value of greater than 83% for the association between subjects not having a sustained viral response to therapy and a genotype comprising HLA-C2 homozygosity combined with at least one LR allele of rs7248668, whereas the positive predictive value of at least one  
25 HLA-C1 allele and at least one of the same LR alleles is only about 62%, and the positive predictive value of two copies of the corresponding high response (HR) allele is only about 60%. The HLA-C2/rs7248668 genotype combination of the invention provides a higher positive predictive value relative to homozygosity of HLA-C2 alleles *solus*, and covers a higher proportion of the population than the IL28B genotype *solus*.

30

Thus, the inventors provide a test for determining the likelihood that a subject will or will not respond to therapy with an immunomodulatory composition, wherein the test provides a high positive predictive value *e.g.*, at least about 70% or 85% or 80% or 81% or 82% or 83% or 84% or 85% or 86% or 87% or 88% or 89% or 90% or 91% or  
5 92% or 93% or 94% or 95% or 96% or 97% or 98% or 99% 100%.

It is to be understood that the test of the present invention may be enhanced or improved by combining any other allele linked to IL28B that is in linkage disequilibrium at least with rs7248668. The inventors have identified one or more  
10 other low response (LR) alleles, each LR allele linked to IL28B, that may be combined with the HLA-C2/rs7248668 genotype combination of the invention *e.g.*, to provide greater population coverage without compromising specificity or positive predictive value of the test. For example, data presented herein demonstrate other HLA-C2/IL28B-LR allele combinations *e.g.*, HLA-C2/rs4803221-LR and/or the HLA-  
15 C2/rs8099917-LR and/or HLA-C2/rs12979860-LR, that each provide positive predictive values of greater than about 75% for subjects not having a sustained viral response to therapy, and which, when combined with the HLA-C2/rs7248668 genotype combination of the invention, may each enhance the value of the test. Such combinations that provide enhanced predictive value the HLA-C2/rs7248668 genotype  
20 combination of the invention are even more desired. For each other HLA-C2/IL28B-LR allele combination to be combined with the HLA-C2/rs7248668-LR genotype, an effect of combining at least one other IL28B-LR allele with HLA-C2 homozygosity is determined to demonstrate that the additional IL28B-LR allele is suitable for combining with the HLA-C2/rs7248668-LR genotype. For example, associations  
25 between the sustained viral response (SVR) of subjects and a genotype comprising HLA-C2 homozygosity and at least one specific IL28B-LR allele selected from rs4803221-LR, rs8099917-LR and rs12979860-LR are provided herein. Thus, referred LR alleles suitable for combining with the HLA-C2/rs7248668 genotype combination of the invention are selected from those set forth in Tables 1, 3-7 hereof *e.g.*, any  
30 IL28B LR allele that is in linkage disequilibrium at least with an LR allele of rs7248668 and/or rs4803221 and/or rs8099917 and/or rs12979860.

By testing other HLA-C2/IL28B-LR genotype combinations, the inventors also provide a test for determining the likelihood that a subject will or will not respond to therapy with an immunomodulatory composition that provides both a high positive predictive value and improved population coverage, *e.g.*, at least about 10% or 11% or 12% or 13% or 14% or 15% or 16% or 17% or 18% or 19% or 20% or 21% or 22% or 23% or 24% or 25% or 26% or 27% or 28% or 29% or 30% or 31% or 32% or 33% or 34% or 35% or 36% or 37% or 38% or 39% or 40% or 41% or 42% or 43% or 44% or 45% or 46% or 47% or 48% or 49% or 50% or 55% or 60% or 65% or 70% or 75% or 80% or 85% or 90% or 95% or 96% or 97% or 98% or 99% of cases.

As used herein, the term "allele linked to IL28B" shall be taken to include any allele positioned within in the 5'-upstream region or 3'-downstream region of the IL28B gene including the IL28A/IL28B intergenic region, or any allele positioned in a promoter region, 3'-untranslated region, exon or intron of the IL28B gene. The allele may be in genomic DNA, and/or in transcribed mRNA where applicable. By way of example, Tables 1, 3-7 hereof provide alleles linked to IL28B that are useful in combination with HLA-C2 homozygosity for determining the likelihood that a subject will respond to treatment with an immunomodulatory composition, especially the likelihood that a subject having a chronic HCV infection will respond to treatment with an immunomodulatory composition comprising IFN- $\alpha$ , such as a composition comprising IFN- $\alpha$  and ribavirin. Preferred alleles linked to IL28B that may be combined with HLA-C2 homozygosity are positioned in the 5'-upstream region or intergenic IL28A/IL28B region of the human genome.

25

For the purposes of nomenclature, the term "rs7248668" shall be taken to mean the polymorphic region that is positioned within the 5'-upstream flanking region of the IL28B gene of humans, wherein said polymorphic region comprises the sequence set forth in SEQ ID NO: 159 or a complementary nucleotide sequence to SEQ ID NO: 159. In the present context, the term "rs7248668" shall also be taken to refer to the single nucleotide polymorphism contained within SEQ ID NO: 159 or a complementary

30

nucleotide sequence to SEQ ID NO: 159. Unless the context requires otherwise, the term "rs7248668 allele" shall be taken to include the HR allele and/or the LR allele of rs7248668.

- 5 The term "HR allele of rs7248668" or "G allele of rs7248668" or "rs7248668-HR" shall be taken to refer to an allele of rs7248668 comprising guanine or guanosine at a position corresponding to position 31 of SEQ ID NO: 160 in the context of a sequence comprising SEQ ID NO: 160, or within the context of a shorter sequence comprising at least about 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or  
10 19 or 20 contiguous nucleotides of SEQ ID NO: 160 including the guanine or guanosine at position 31 of SEQ ID NO: 160. Functionally, the HR allele of rs7248668 is associated with a better patient response in the population to antiviral therapy with an immunomodulatory agent such as pegylated IFN- $\alpha$  or a combination comprising pegylated IFN- $\alpha$  and ribavirin, and preferably a sustained virological response (SVR)  
15 to said therapy.

- The term "LR allele of rs7248668" or "A allele of rs7248668" or "rs7248668-LR" shall be taken to refer to an allele of rs7248668 comprising adenine or adenosine at a position corresponding to position 31 of SEQ ID NO: 161 in the context of a sequence  
20 comprising SEQ ID NO: 161, or within the context of a shorter sequence comprising at least about 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 contiguous nucleotides of SEQ ID NO: 161 including the adenine or adenosine at position 31 of SEQ ID NO: 161. Functionally, the LR allele of rs7248668 is associated with a weak or poor patient response in the population to antiviral therapy  
25 with an immunomodulatory agent such as pegylated IFN- $\alpha$ , or a combination comprising pegylated IFN- $\alpha$  and ribavirin.

- The term "HLA-C2 allele" shall be taken to mean an allele of the HLA-C gene of humans that encodes an HLA-C polypeptide comprising a lysine (K) residue at a  
30 position corresponding to position 80 of a full-length HLA-C polypeptide. Particularly preferred HLA-C2 molecules will also comprise an asparagine (N) residue at a position

corresponding to position 77 of a full-length HLA-C polypeptide. Functionally, an HLA-C2 allele is associated with a weak or poor patient response in the population to antiviral therapy with an immunomodulatory agent such as pegylated IFN- $\alpha$  or a combination comprising pegylated IFN- $\alpha$  and ribavirin. Exemplary HLA-C2  
5 molecules are selected from HLA-Cw2, HLA-Cw4, HLA-Cw5, HLA-Cw6, HLA-Cw15, HLA-Cw17, HLA-Cw18, HLA-Cw\*0114, HLA-Cw\*0307, HLA-Cw\*0315, HLA-Cw\*0345, HLA-Cw\*0707, HLA-Cw\*0709, HLA-Cw\*0776, HLA-Cw\*0810, HLA-Cw\*1204, HLA-Cw\*1205, HLA-Cw\*1209, HLA-Cw\*1221, HLA-Cw\*1233, HLA-Cw\*1241, HLA-Cw\*1412, HLA-Cw\*1602, HLA-Cw\*1609, HLA-Cw\*1612,  
10 HLA-Cw\*1619, and HLA-Cw\*1625, and being other than HLA-Cw\*0227, HLA-Cw\*0411, HLA-Cw\*0429, HLA-Cw\*0436, HLA-Cw\*0455, HLA-Cw\*0611, HLA-Cw\*1507, HLA-Cw\*1525, or HLA-Cw\*1543. The amino acid sequences of HLA-C2 alleles are publicly-available.

15 The term "HLA-C2 homozygosity" or similar term refers to two copies of any HLA-C2 allele at the same locus in the genome of a subject.

The term "HLA-C1 allele" shall be taken to mean an allele of the HLA-C gene of humans that encodes an HLA-C polypeptide comprising an asparagine (N) residue at a  
20 position corresponding to position 80 of a full-length HLA-C polypeptide. Particularly preferred HLA-C1 molecules will also comprise a serine (S) residue at a position corresponding to position 77 of a full-length HLA-C polypeptide. Functionally, an HLA-C1 allele is associated with a better patient response in the population to antiviral therapy with an immunomodulatory agent such as pegylated IFN- $\alpha$  or a combination  
25 comprising pegylated IFN- $\alpha$  and ribavirin, and preferably a sustained virological response (SVR) to said therapy. Exemplary HLA-C1 molecules include HLA-Cw1, HLA-Cw3, HLA-Cw7, HLA-Cw8, HLA-Cw12, HLA-Cw13, HLA-Cw14, HLA-Cw16, HLA-Cw\*0227, HLA-Cw\*0411, HLA-Cw\*0429, HLA-Cw\*0436, HLA-Cw\*0455, HLA-Cw\*0611, HLA-Cw\*1507, HLA-Cw\*1525, and HLA-Cw\*1543, and being other  
30 than HLA-Cw\*0114, HLA-Cw\*0307, HLA-Cw\*0315, HLA-Cw\*0345, HLA-Cw\*0707, HLA-Cw\*0709, HLA-Cw\*0776, HLA-Cw\*0810, HLA-Cw\*1204, HLA-

Cw\*1205, HLA-Cw\*1209, HLA-Cw\*1221, HLA-Cw\*1233, HLA-Cw\*1241, HLA-Cw\*1412, HLA-Cw\*1602, HLA-Cw\*1609, HLA-Cw\*1612, HLA-Cw\*1619, or HLA-Cw\*1625. The amino acid sequences of HLA-C1 alleles are publicly-available.

- 5 For the purposes of nomenclature, the term "rs4803221" shall be taken to mean the polymorphic region that is positioned within the 5'-upstream flanking region of the IL28B gene of humans, wherein said polymorphic region comprises the sequence set forth in SEQ ID NO: 41 or a complementary nucleotide sequence thereto. In the present context, the term "rs4803221" shall also be taken to refer to the single  
10 nucleotide polymorphism contained within SEQ ID NO: 41 *i.e.*, at position 31 thereof. Unless the context requires otherwise, the term "rs4803221" shall be taken to refer to the HR allele and/or the LR allele of rs4803221.

The term "HR allele of rs4803221" or "C allele of rs4803221" or "rs4803221-HR"  
15 shall be taken to refer to an allele of rs4803221 comprising cytosine at a position corresponding to position 31 of SEQ ID NO: 162 in the context of a sequence comprising SEQ ID NO: 162, or within the context of a shorter sequence comprising at least about 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 contiguous nucleotides of SEQ ID NO: 162 including the cytosine at position  
20 31 of SEQ ID NO: 162. Functionally, the HR allele of rs4803221 is associated with a better patient response in the population to antiviral therapy with an immunomodulatory agent such as pegylated IFN- $\alpha$  or a combination comprising pegylated IFN- $\alpha$  and ribavirin, and preferably a sustained virological response (SVR) to said therapy.

25

The term "LR allele of rs4803221" or "G allele of rs4803221" or "rs4803221-LR" shall be taken to refer to an allele of rs4803221 comprising guanine or guanosine at a position corresponding to position 31 of SEQ ID NO: 163 in the context of a sequence comprising SEQ ID NO: 163, or within the context of a shorter sequence comprising at  
30 least about 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 contiguous nucleotides of SEQ ID NO: 163 including the guanine or

guanosine at position 31 of SEQ ID NO: 163. Functionally, the LR allele of rs4803221 is associated with a weak or poor patient response in the population to antiviral therapy with an immunomodulatory agent such as pegylated IFN- $\alpha$ , or a combination comprising pegylated IFN- $\alpha$  and ribavirin.

- 5
- For the purposes of nomenclature, the term "rs12979860" shall be taken to mean the polymorphic region that is positioned within the 5'-upstream flanking region of the IL28B gene of humans, wherein said polymorphic region comprises the sequence set forth in SEQ ID NO: 42 or a complementary nucleotide sequence thereto. In the
- 10 present context, the term "rs12979860" shall also be taken to refer to the single nucleotide polymorphism contained within SEQ ID NO: 42 *i.e.*, at position 31 thereof. Unless the context requires otherwise, the term "rs12979860" shall be taken to refer to the HR allele and/or the LR allele of rs12979860.
- 15 The term "HR allele of rs12979860" or "C allele of rs12979860" or "rs12979860-HR" shall be taken to refer to an allele of rs12979860 comprising cytosine at a position corresponding to position 31 of SEQ ID NO: 164 in the context of a sequence comprising SEQ ID NO: 164, or within the context of a shorter sequence comprising at least about 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or
- 20 19 or 20 contiguous nucleotides of SEQ ID NO: 164 including the cytosine at position 31 of SEQ ID NO: 164. Functionally, the HR allele of rs12979860 is associated with a better patient response in the population to antiviral therapy with an immunomodulatory agent such as pegylated IFN- $\alpha$  or a combination comprising pegylated IFN- $\alpha$  and ribavirin, and preferably a sustained virological response (SVR)
- 25 to said therapy.

- The term "LR allele of rs12979860" or "T" allele of rs12979860" or "rs12979860-LR" shall be taken to refer to an allele of rs12979860 comprising thymidine or uracil at a position corresponding to position 31 of SEQ ID NO: 165 in the context of a sequence
- 30 comprising SEQ ID NO: 165, or within the context of a shorter sequence comprising at least about 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or

19 or 20 contiguous nucleotides of SEQ ID NO: 165 including the thymidine or uracil at position 31 of SEQ ID NO: 162. Functionally, the LR allele of rs12979860 is associated with a weak or poor patient response in the population to antiviral therapy with an immunomodulatory agent such as pegylated IFN- $\alpha$ , or a combination  
5 comprising pegylated IFN- $\alpha$  and ribavirin.

The term "rs8099917" shall be taken to mean the polymorphic region that is positioned within the 5'-upstream flanking region of the IL28B gene of humans, wherein said polymorphic region comprises the sequence set forth in SEQ ID NO: 4 or a  
10 complementary nucleotide sequence thereto. In the present context, the term "rs8099917" shall also be taken to refer to the single nucleotide polymorphism contained within SEQ ID NO: 4 *i.e.*, at position 31 thereof. Unless the context requires otherwise, the term "rs8099917" shall be taken to refer to the HR allele and/or the LR allele of rs8099917.

15

The term "HR allele of rs8099917" or "T allele of rs8099917" or "rs8099917-HR" shall be taken to refer to an allele of rs8099917 comprising thymidine or uracil at a position corresponding to position 31 of SEQ ID NO: 5 in the context of a sequence comprising SEQ ID NO: 5, or within the context of a shorter sequence comprising at  
20 least about 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 contiguous nucleotides of SEQ ID NO: 5 including the thymidine (or a uracil) at position 31 of SEQ ID NO: 5. Functionally, the HR allele of rs8099917 is associated with a better patient response in the population to antiviral therapy with an immunomodulatory agent such as pegylated IFN- $\alpha$  or a combination comprising  
25 pegylated IFN- $\alpha$  and ribavirin, and preferably a sustained virological response (SVR) to said therapy.

The term "LR allele of rs8099917" or "G allele of rs8099917" or "rs809917-LR" shall be taken to refer to an allele of rs8099917 comprising guanine or guanosine at a  
30 position corresponding to position 31 of SEQ ID NO: 6 in the context of a sequence comprising SEQ ID NO: 6, or within the context of a shorter sequence comprising at

least about 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 contiguous nucleotides of SEQ ID NO: 6 including the guanine or guanosine at position 31 of SEQ ID NO: 6. Functionally, the LR allele of rs8099917 is associated with a weak or poor patient response in the population to antiviral therapy with an immunomodulatory agent such as pegylated IFN- $\alpha$ , or a combination comprising pegylated IFN- $\alpha$  and ribavirin.

Other exemplary alleles linked to IL28B that may be combined with the rs7248668/HLA-C genotype combination will preferably be selected from those listed in the accompanying Tables 1 and 3-7 hereof, which each have a high significance in their association with high response (HR) or low response (LR) to therapy.

Thus, the data herein provide the basis for prognostic tests to determine the likelihood that a subject will or will not respond to therapy with one or more immunomodulatory agents, and in particular, the likelihood that a subject suffering from a chronic HCV infection will or will not respond to therapy comprising immunomodulatory agents such as IFN- $\alpha$  and ribavirin. The tests herein based on homozygosity of HLA-C2 alleles in combination with at least one rs7248668-LR allele have improved accuracy compared to tests based on a single HLA-C2 allele or a single HLA-C1 allele in combination with at least one rs7248668-HR allele or two rs7248668-HR alleles in determining the likelihood of a subject's response to therapy. Additionally, the data herein support multi-analyte prognostic tests for homozygosity of HLA-C2 alleles in combination with at least two LR alleles at different loci and each linked to IL28B, wherein one of said LR alleles is rs7248668-LR. Such multi-analyte tests combine the improved accuracy and specificity of the HLA-C2/rs7248668 genotype combination of the invention with greater population coverage. For example, data presented in Table 12 and Table 13 provide support for extrapolating tests for the HLA-C2/rs7248668 genotype combination to include additional IL28B-LR alleles *e.g.*, at least one LR allele rs4803221 and/or at least one LR allele of rs8099917 and/or at least one LR allele of rs12979680, with PPV of more than 80% and a population coverage of up to about 60%.

Accordingly, the HLA-C2/rs7248668 genotype combination, optionally with one or more additional IL28B alleles, serotypes, haplotypes and genotypes provided herein, provide the means for accurately determining the likelihood that a subject will or will  
5 not respond to therapy comprising an immunomodulatory composition.

As used herein, the terms “accurately determining” or “accurate prognosis” shall be taken to mean an association of a particular allele, genotype, serotype, genotype combination or haplotype with a high response (HR) or low response (LR) to therapy,  
10 or an association of a particular allele, genotype, serotype, genotype combination or haplotype with a non-response to therapy, or an association of a particular allele, genotype, serotype, genotype combination or haplotype with relapse. Accurate determination or prognosis may be statistically significant, *e.g.*, at  $p < 10^{-3}$ , or preferably  $p < 10^{-4}$ , or more preferably  $p < 10^{-5}$  or  $p < 10^{-6}$  or  $p < 10^{-7}$ . Alternatively, or  
15 in addition, the term “accurately determining” means that a plurality of alleles, genotypes, serotypes, a genotype combination or a haplotype, *e.g.*, HLA-C2 in combination with at least one rs7248668-LR allele, or optionally in combination with at least one additional IL28B-LR allele, provides a positive predictive value (PPV) with respect to treatment outcome of more than about 75% for patients in a population. The  
20 PPV may increase proportionately with the number of markers assayed. Preferably, an association is stronger than for rs8099917 *solus* in a population. The PPV may increase proportionately with the number of markers assayed.

In this context, the term “population” means a test population of greater than 100  
25 matched individuals or greater than 200 matched individuals or greater than 300 matched individuals or greater than 400 matched individuals or greater than 500 matched individuals.

By “matched” is meant that the individuals of the test population have similar or near-  
30 identical age, BMI, viral titer, and treatment regime. For practical purposes, the present invention also provides for accurate prognosis in a “real world” population of

individuals suffering from the same medical condition *e.g.*, individuals suffering from the same condition that are at least matched with respect to ethnicity. By way of explanation and without limitation, one example of the invention provides for accurate prognosis of treatment for primary or chronic HCV infection in a population of  
5 Caucasian patients.

As used herein, the term “immunomodulatory composition” shall be taken in its broadest context to mean a composition comprising one or more compounds capable of modulating expression or secretion of one or more cytokines involved in autoimmunity  
10 and/or immune responses to infectious agents, or by modulating one or more components of a cytokine signalling pathway. The term “compound” in this context includes a protein, small molecule, antibody molecule, or nucleic acid *e.g.*, RNAi, antisense RNA, ribozyme or siRNA.

15 Thus, a HLA-C2/rs7248668 genotype combination of the present invention, optionally with one or more IL28B additional alleles, serotypes, haplotypes, or genotypes provided herein, clear application for the accurate prognosis of a response to any therapy comprising administration of an “immunomodulatory composition” that is known to be used and/or known to be useful in the treatment of a viral infection and/or  
20 neoplasia and/or Th1-mediated disease and/or Th2-mediated disease.

For example, the invention is suitable for accurate prognosis of a response to therapy comprising administration of an “immunomodulatory composition” for treatment of Th1-mediated disease and/or Th2-mediated disease *e.g.*, one or more conditions  
25 selected individually or collectively from the group consisting of multiple sclerosis (MS), rheumatoid arthritis (RA), Type I diabetes (IDDM), scleroderma, Con A hepatitis, atopic dermatitis, asthma, allergic rhinitis and allergy. Alternatively, or in addition, the invention is suitable for accurate prognosis of a response to therapy comprising administration of an “immunomodulatory composition” for treatment of  
30 one or more infections by viruses selected individually or collectively from the group consisting of human papillomaviruses (*e.g.*, papillomavirus(es) selected from HPV16,

HPV6 and HPV11), herpes viruses (*e.g.*, herpes virus(es) selected from HSV-1, HSV-2, VZV, HHV-6, HHV-7, HHV-8 (KSHV), HCMV and EBV), picornaviruses (*e.g.*, picornavirus(es) selected from Coxsackie B virus(es) and EMCV), flaviviruses (*e.g.*, flavivirus(es) selected from encephalitis virus(es) and hepatitis virus(es) such as HAV and/or HBV and/or HCV), arenaviruses (arenavirus(es) associated with a viral haemorrhagic fever); togaviruses (togavirus(es) selected from equine encephalitis viruses), bunyaviruses (*e.g.*, bunyavirus(es) selected from Rift Valley fever virus, Crimean-Congo haemorrhagic fever virus, HTNV and APEUV), filoviruses (*e.g.*, filovirus(es) selected from Ebola virus and Marburg virus), paramyxoviruses (*e.g.*, RSV), rhabdoviruses (*e.g.*, VSV), orthomyxoviruses (*e.g.*, influenza viruses such as IAV), and coronaviruses (*e.g.*, SARS-associated coronavirus, "SARS-CoV"). For example, the invention provides means for prognosis of a response to therapy comprising administration of an "immunomodulatory composition" for treatment of one or more infections by hepatitis virus(es), such as HAV and/or HBV and/or HCV, and especially HCV. Alternatively, or in addition, the invention is suitable for accurate prognosis of a response to therapy comprising administration of an "immunomodulatory composition" for treatment of one or more neoplasias or pre-cancerous conditions, such as neoplasia(s) and pre-cancerous condition(s) selected individually or collectively from the group consisting of HPV-associated cancer (*e.g.*, cervical intraepithelial neoplasia and/or cervical carcinoma and/or vulvar intraepithelial neoplasia and/or penile intraepithelial neoplasia and/or perianal intraepithelial neoplasia), hepatocellular carcinoma, basal cell carcinoma, squamous cell carcinoma, actinic keratosis, melanoma, hairy cell leukemia, Kaposi's sarcoma, non-Hodgkin's lymphoma, astrocytoma, glioblastoma, thymoma, fibrosarcoma.

25

In another example, a HLA-C2/rs7248668 genotype combination of the present invention, optionally with one or more IL28B additional alleles, serotypes, haplotypes, or genotypes provided herein, provides the means for accurately determining the likelihood that a subject will respond to therapy comprising of an immunomodulatory composition comprising IFN.

30

Unless the context requires otherwise, the term "IFN" as used herein shall be taken to include any known interferon molecule *e.g.*, IFN- $\alpha$ , IFN- $\beta$ , IFN- $\omega$ , IFN- $\gamma$ , IFN- $\lambda$ 1, IFN- $\lambda$ 2, or IFN- $\lambda$ 3, a composition comprising a plurality of any interferon molecules *e.g.*, two or more molecules selected from IFN- $\alpha$ , IFN- $\beta$ , IFN- $\omega$ , IFN- $\gamma$ , IFN- $\lambda$ 1, IFN- $\lambda$ 2 and  
5 IFN- $\lambda$ 3, a composition comprising one or more derivatives of an interferon molecule *e.g.*, a pegylated interferon, and mixtures of said one or more derivatives with one or more non-derivative interferon molecules.

For example, the present invention has clear application for the prognosis of a response  
10 to any therapy comprising administration of "IFN" that is known to be used and/or known to be useful in the treatment of a viral infection and/or neoplasia and/or Th1-mediated disease and/or Th2-mediated disease. For example, the invention is useful for prognosis of a response to an infection treatable by "IFN", wherein the infection is by one or more ssRNA viruses, *i.e.*, an infection by one or more (+) ssRNA viruses and/or  
15 an infection by one or more (-)ssRNA viruses, such as SARS-associated coronavirus (SARS-CoV), HBV, HCV, coxsackie B virus, EMCV, Ebola virus, VSV, IAV, HTNV, or APEUV, and/or one or more double-stranded DNA viruses such as HSV-1 or HSV-2. Alternatively, or in addition, the invention is useful for prognosis of a pre-cancerous lesion or neoplasia treatable by "IFN" *e.g.*, a pre-cancerous lesion or neoplasia selected  
20 from the group consisting of condylomata acuminata, hairy cell leukemia, Kaposi's sarcoma, melanoma, non-Hodgkin's lymphoma, astrocytoma, glioblastoma, thymoma and fibrosarcoma. Alternatively, or in addition, the invention is useful for prognosis of a Th1-mediated disease or Th2-mediated disease treatable by "IFN" *e.g.*, a disease selected from the group consisting of MS, asthma and Con A-induced hepatitis.

25

In another example, a HLA-C2/rs7248668 genotype combination of the present invention, optionally with one or more IL28B additional alleles, serotypes, haplotypes, or genotypes provided herein, provides the means for accurately determining the likelihood that a subject will respond to therapy comprising an immunomodulatory  
30 composition comprising guanosine analog(s).

Unless the context requires otherwise, the term "guanosine analog" as used herein shall be taken to include any known guanosine analog, a composition comprising a plurality of guanosine analogs, a composition comprising one or more derivatives of one or more guanosine analogs and mixtures of said one or more derivatives with one or more  
5 non-derivative guanosine analogs. Preferred guanosine analogs in this context are those compounds that are capable of modulating levels of Th1 and/or Th2 cells, or that have antiviral and/or anti-cancer activity. Exemplary guanosine analogs are selected from ribavirin, viramidine, 7-benzyl-8-bromoguanine, 9-benzyl-8-bromoguanine, and CpG-containing oligonucleotide(s), and derivative(s), salt(s), solvate(s) and hydrate(s)  
10 thereof *e.g.*, ribavirin 5'-monophosphate, ribavirin 5'-diphosphate, ribavirin 5'-triphosphate, and ribavirin 3',5'-cyclic phosphate.

In another example, a HLA-C2/rs7248668 genotype combination of the present invention, optionally with one or more IL28B additional alleles, serotypes, haplotypes,  
15 or genotypes provided herein, provides the means for accurately determining the likelihood that a subject will respond to therapy comprising an immunomodulatory composition comprising IFN and guanosine analog(s).

This is the first disclosure of interactions *e.g.*, additive interaction or epistatic  
20 interaction between HLA-C2 alleles and one or more LR alleles of rs7248668 in determining treatment outcome. In one example, the inventors found that homozygosity for the HLA-C2 allele in combination with one or two copies of a low response (LR) allele of rs7248668 linked to IL28B is predictive of a poor response to antiviral therapy. The present invention thus provides the first disclosure of an  
25 association between a genotype comprising two HLA-C2 alleles and at least one rs7248668-LR allele, and poor treatment outcome.

The alleles, serotypes, haplotypes and genotypes disclosed in Tables 1, 3-7, 12 and 13 hereof comprise allelic variants and combinations of allelic variants that are associated  
30 with a high response (HR) to therapy or a low response (LR) to therapy. Accordingly, the present invention clearly encompasses the use of the HLA-C2/rs7248668 genotype

- combination of the present invention and any additional IL28B-LR allele(s) and/or any other IL28B-HR allele(s) disclosed in Tables 1, 3-7, 12 or 13 hereof for determining the likelihood that a subject will respond to therapy comprising an immunomodulatory composition as described herein. The only requirement to such combinations is that they provide the same or better population coverage without substantial loss in predictive value of the HLA-C2/rs7248668 genotype combination and/or without substantial reduction in assay specificity, and/or provide enhanced predictive value relative to the HLA-C2/rs7248668 genotype combination.
- 10 It will be apparent to the skilled artisan that any gene or gene fragment may be employed in performing the present invention by virtue of the gene or gene fragment being capable of detecting alleles of HLA-C and rs7248668 disclosed herein, and optionally any additional allele of IL28B disclosed herein. An allele of HLA-C not specifically disclosed herein may also provide a surrogate means of detecting an HLA-C allele disclosed herein by virtue of being in linkage disequilibrium with it. An allele of IL28B not specifically disclosed herein may also provide a surrogate means of detecting an IL28B allele disclosed herein, other than rs7248668, by virtue of being in linkage disequilibrium with it. For example, a gene or gene fragment will be capable of detecting an allele of HLA-C or IL28B if it hybridizes to an allele *e.g.*, a SNP, microsatellite or INDEL in linkage disequilibrium with an HLA-C or IL28B allele disclosed herein. Similarly, a gene or gene fragment hybridising to one allele of a haplotype block may provide a surrogate for detection of the haplotype block by virtue of the alleles of the haplotype block being in linkage disequilibrium. Preferred markers that may provide surrogates for an exemplified allele hereof are generally positioned within 5kb of the 5'-end or 3'-end of the gene, the intergenic IL28A/IL28B region, or with the protein-encoding region or an intron region of the structural HLA-C or IL28B gene, or within the protein-encoding region of the HLA-C gene.

For example, the haplotype block identified and characterized by the inventors for the IFN- $\lambda$ 3 gene (Table 6), and expression data (Figure 1) demonstrating that expression of IFN- $\lambda$ 2 and IFN- $\lambda$ 3 is reduced in carriers of the LR allele *i.e.*, the G allele, of rs

8099917 relative to carriers of the corresponding HR allele *i.e.*, the T allele, indicate that all of the chromosome 19 SNPs presented in Table 1 are definitely linked to the IFN- $\lambda$ 3 gene, with the possible exception of rs4803224, rs12980602 and rs10853728. The excluded SNPs under these criteria are more distal than rs8099917 from the structural gene region *i.e.*, encoding IFN- $\lambda$ 3. Thus, the present invention may be employed by identifying any marker linked to HLA-C2 or IL28B and associated with treatment outcome *e.g.*, in the 5'-upstream region or an intron or an exon or the 3'-downstream region.

Accordingly, the present invention encompasses, but is not limited to, the use of HLA-C2 homozygosity combined with at least one rs7248668-LR allele, optionally in combination with any other IL28B allele, preferably any other IL28B-LR allele, set forth in any one or more of Tables 1, 3-7, and any combination thereof *e.g.*, a specific haplotype set forth in any one of Tables 3-6, for determining the likelihood that a subject will or will not respond to therapy comprising an immunomodulatory composition as described herein. For example, markers linked to the HLA-C2 or rs7248668 genes or fragments thereof, optionally in combination with one or more other markers linked to one or more other IL28B allelic variant genes or a fragment thereof, may be employed for accurate prognosis. By "fragment" in this context, is meant a portion of a gene of sufficient length to be useful for detection of gene expression associated with the polymorphism and/or of sufficient length to directly identify the polymorphism *e.g.*, in a platform suitable for identifying SNPs as described herein.

## 2. *Specific embodiments*

The scope of the invention will be apparent from the claims as filed with the application that follow the examples. The claims as filed with the application are hereby incorporated into the description. The scope of the invention will also be apparent from the following description of specific embodiments.

In one example, the present invention provides a method for accurately determining the likelihood that a subject will or will not respond to treatment with an immunomodulatory composition, said method comprising detecting homozygosity of a HLA-C2 allele and at least one IL28B low response (LR) allele in a sample from the  
5 subject, wherein the IL28B LR allele is the LR allele of rs7248668, and wherein detection of said homozygosity and said at least one LR allele is indicative of a low response of the subject to treatment with said composition.

In another example, the present invention provides a method for accurately determining  
10 the likelihood that a subject will or will not respond to treatment with an immunomodulatory composition, said method comprising detecting homozygosity of HLA-C2 and rs7248668-LR alleles in a sample from the subject, wherein detection of the double homozygosity is indicative of a low response of the subject to treatment with said composition.

15

In another example, the present invention provides a method for accurately determining the likelihood that a subject will or will not respond to treatment with an immunomodulatory composition, said method comprising detecting two or more markers in a sample from the subject, wherein at least one marker is linked to an HLA-  
20 C2 allele and at least one marker is linked to the low response (LR) allele of rs7248668, and wherein said detection is indicative of HLA-C2 homozygosity and at least one LR allele of rs7248668 in the genome of the subject, said indication being further indicative of a low response of the subject to treatment with said composition.

25 In another example, the present invention provides a method for accurately determining the likelihood that a subject will or will not respond to treatment with an immunomodulatory composition, said method comprising performing a method according to any example *supra*, and further comprising detecting one or more low response (LR) alleles of IL28B in a sample from the subject, wherein detection of  
30 HLA-C2 homozygosity and at least one rs7248668-LR allele and at least one additional

LR allele of IL28B is indicative of a low response of the subject to treatment with said composition.

The rs7248668-LR allele is defined herein. For example, the rs7248668-LR allele may  
5 comprise a adenine or adenosine at a position corresponding to position 31 of SEQ ID  
NO: 161 in the context of said sequence or a shorter sequence comprising at least about  
5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20  
contiguous nucleotides of SEQ ID NO: 161 including the adenine or adenosine at  
position 31 thereof, a sequence complementary thereto.

10

Exemplary HLA-C2 alleles are selected from HLA-Cw2, HLA-Cw4, HLA-Cw5, HLA-  
Cw6, HLA-Cw15, HLA-Cw17, HLA-Cw18, HLA-Cw\*0114, HLA-Cw\*0307, HLA-  
Cw\*0315, HLA-Cw\*0345, HLA-Cw\*0707, HLA-Cw\*0709, HLA-Cw\*0776, HLA-  
Cw\*0810, HLA-Cw\*1204, HLA-Cw\*1205, HLA-Cw\*1209, HLA-Cw\*1221, HLA-  
15 Cw\*1233, HLA-Cw\*1241, HLA-Cw\*1412, HLA-Cw\*1602, HLA-Cw\*1609, HLA-  
Cw\*1612, HLA-Cw\*1619, and HLA-Cw\*1625, and being other than HLA-Cw\*0227,  
HLA-Cw\*0411, HLA-Cw\*0429, HLA-Cw\*0436, HLA-Cw\*0455, HLA-Cw\*0611,  
HLA-Cw\*1507, HLA-Cw\*1525, or HLA-Cw\*1543. The amino acid sequences of  
HLA-C2 alleles are publicly-available.

20

Exemplary additional low response (LR) alleles of IL28B are selected from the group  
of single nuclear polymorphisms (SNPs) set forth in any one of Table 1 or comprise a  
SNP set forth in Table 1 or are encoded by nucleic acid comprising a SNP set forth in  
Table 1 or linked to a SNP set forth in Table 1.

25

An additional IL28B-LR allele may consist of a SNP set forth in Table 1, or may  
consist of a SNP set forth in Table 3 or comprise a SNP set forth in Table 3 or be  
encoded by nucleic acid comprising a SNP set forth in Table 3 or be linked to a SNP  
set forth in Table 3. In another example, an additional IL28B-LR allele may consist of  
30 a SNP set forth in Table 4 or 5 or comprise a SNP set forth in Table 4 or 5 or be  
encoded by nucleic acid comprising a SNP set forth in Table 4 or 5 or be linked to a

SNP set forth in Table 4 or 5. Alternatively, or in addition, an additional IL28B-LR allele may comprise a sequence selected from the group consisting of:

- (i) a sequence of an IL28B-LR allele set forth in any one of SEQ ID NOs: 1 to 60, 62, 64 to 67, 69, 71 to 74, 76, 78, 79, 81, 83 to 157, or 162-165, and preferably being  
5 selected from SEQ ID NOs: 6, 11, 69, 86, 89, 92, 95, 98, 100, 103, 106, 110, 112, 115, 119, 122, 124, 127, 131, 134, 137, 140, 142, 145, 149, 151, 154, 157, 161, 163 and 165, and more preferably being selected from SEQ ID Nos: 6, 11, 69, 86, 89, 163 and 165; and
- (ii) a sequence complementary to a sequence at (i).

10

In another example, an additional IL28B-LR allele may be the rs4803221-LR allele *e.g.*, SEQ ID NO: 163 or at least about 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 contiguous nucleotides thereof including the nucleotide at position 31 thereof, or a complementary sequence thereto.

15

In another example, an additional IL28B-LR allele may be the rs8099917-LR allele *e.g.*, SEQ ID NO: 6 or at least about 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 contiguous nucleotides thereof including the nucleotide at position 31 thereof, or a complementary sequence thereto.

20

In another example, an additional IL28B-LR allele may be the rs12979860-LR allele *e.g.*, SEQ ID NO: 165 or at least about 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 contiguous nucleotides thereof including the nucleotide at position 31 thereof, or a complementary sequence thereto.

25

It is clearly within the scope of the invention to detect two or more additional IL28B-LR alleles at different loci of IL28B, such as a haplotype comprising two or more IL28B alleles *e.g.*, wherein the haplotype comprises at least one or two or three alleles of rs4803221, rs12979860, or rs8099917, such as a haplotype comprising alleles of  
30 rs12980275, rs8105790, rs8103142, rs10853727, rs8109886 and rs8099917.

In the foregoing examples, one or two LR alleles at any additional locus of IL28B set forth in Tables 1 and 3-7 may be detected. Thus, heterozygosity or homozygosity at any additional locus may be employed in combination with the HLA-C2/rs7248668-LR genotype combination of the invention.

5

It is clearly within the scope of the invention as described according to any example hereof to detect a plurality of IL28B genotypes or IL28B markers *e.g.*, two or three or four or five or six or seven or more IL28B genotypes or markers, in addition to the HLA-C2/rs7248668-LR genotype combination of the invention.

10

The present invention also encompasses the detection of a modified level of expression *e.g.*, increased or reduced expression of HLA-C and/or IL28B in a sample from the subject, wherein said modified expression is indicative of a response of the subject to treatment with said composition. Modified expression may be determined by *e.g.*,  
15 increased or reduced expression of one or more of the genes or proteins carrying a diagnostic polymorphism. To detect modified expression, a modified level of at least one expression product of the gene(s) is detected *e.g.*, by nucleic acid-based assay or antigen-based assay. For example, an amplification reaction, *e.g.*, isothermal amplification or PCR reaction such as RT-PCR, is performed to detect an mRNA  
20 transcript of the gene(s) in a sample from the subject. Alternatively, to detect expressed protein, a protein-containing sample derived from a subject is contacted with an antibody or ligand capable of specifically binding to an allelic variant of a protein encoded by the gene(s) said marker for a time and under conditions sufficient for complex to form and the complex is detected. Any standard immunoassay may be  
25 employed *e.g.*, ELISA, including sandwich ELISA performed in a microtiter well or in a lateral flow or flow-through assay format. In any assay to determine expression, it is possible to control for variability *e.g.*, by comparing expression in the sample to expression in a control sample. Preferred control samples are selected from the group consisting of: (i) sample(s) from one or more subjects not being treated with the  
30 immunomodulatory composition; and (ii) a data set comprising measurements of expression determined previously for the sample(s) at (i).

In performing the prognostic method of the invention, or any diagnostic or therapeutic assay or process employing the method, the sample will generally comprise genomic DNA, mRNA, protein or a derivative thereof. Amplified DNA or cDNA derived from  
5 genomic DNA or mRNA is also useful. Accordingly, a nucleated cell and/or an extract thereof comprising protein or nucleic acid, is particularly useful if the assay is nucleic acid-based or protein-based. For protein-based assays *e.g.*, immunoassay, the sample should comprise cell extract expected to comprise the marker protein *e.g.*, a cell expressing IL28B. Accordingly, the present invention encompasses the use of any  
10 sample selected from the group consisting of whole blood, serum, plasma, peripheral blood mononuclear cells (PBMC), a buffy coat fraction, saliva, urine, a buccal cell, liver biopsy and a skin cell or combinations thereof.

It is to be understood that the invention may be performed *ex vivo i.e.*, wherein the  
15 sample has been derived or isolated or obtained previously from the subject.

The sample may comprise genomic DNA, mRNA, protein or a derived thereof.

In accordance with the prognostic method of the invention as described according to  
20 any example hereof, a positive response may be selected from the group consisting of:  
(i) a response comprising enhanced clearance of a virus or a reduction in virus titer or a change in other health characteristic of the subject related to reduced virus titer or enhanced clearance; (ii) a response comprising recovery or remission from cancer or reduced growth of a tumor or pre-cancerous lesion; (iii) a change in Th1 cell number,  
25 Th2 cell number or Th1/Th2 cell balance or a change in other health characteristic of the subject indicative of recovery from a Th1-mediated or Th2-mediated disease; and  
(iv) a combination of two or all of (i) to (iii).

Similarly, a low response or non-response may be selected from the group consisting  
30 of: (i) a failure to clear of a virus or to reduce virus titer or change in other health characteristic of the subject related to said failure; (ii) a failure to recover or enter

remission from cancer or to reduce growth of a tumor or pre-cancerous lesion; (iii) no significant change in Th1 cell number, Th2 cell number or Th1/Th2 cell balance or health characteristic of the subject that may indicate recovery from a Th1-mediated or Th2-mediated disease; and (iv) a combination of two or all of (i) to (iii).

5

For those diseases and conditions in which racial origin or genetic background is significant in the association with response, it is preferred that the subject belongs to that racial background or has a matching genetic background. In one example, the subject is Caucasian *e.g.*, northern European. Alternatively, the subject may be African *e.g.*, Zulu, or Asian *e.g.*, Chinese.

10

The immunomodulatory composition may also comprise one or more IFNs and/or one or more derivatives of said one or more of said IFNs *e.g.*, one or more IFNs selected from IFN- $\alpha$ , IFN- $\beta$ , IFN- $\omega$ , IFN- $\gamma$ , IFN- $\lambda$ 1, IFN- $\lambda$ 2 and IFN- $\lambda$ 3 and/or one or more derivatives of any one or more of said IFNs. Alternatively, or in addition the immunomodulatory composition may comprise one or more guanosine analogs and/or one or more derivatives of said one or more of said guanosine analogs *e.g.*, one or more of ribavirin, viramidine, 7-benzyl-8-bromoguanine, 9-benzyl-8-bromoguanine, and CpG-containing oligonucleotide(s), and derivative(s), salt(s), solvate(s) and hydrate(s) thereof. For example, the immunomodulatory composition comprises IFN- $\alpha$  and ribavirin. Testing of responses to pegylated IFNs are clearly encompassed.

15

20

In another example, the present invention provides a process for accurately determining the likelihood that a subject will or will not respond to treatment of Th1-mediated disease and/or Th2-mediated disease with an immunomodulatory composition, said process comprising performing a method according to any example hereof for accurately determining a likely response of a subject to treatment with an immunomodulatory composition, and determining a response for the subject selected from the group consisting of:

25

(i) a change in Th1 cell number, Th2 cell number or Th1/Th2 cell balance or a change in other health characteristic of the subject indicative of recovery from a Th1-

30

mediated or Th2-mediated disease, wherein said response is indicative of a response to treatment; and

- (ii) no significant change in Th1 cell number, Th2 cell number or Th1/Th2 cell balance or health characteristic of the subject that may indicate recovery from a Th1-mediated or Th2-mediated disease, wherein said response is indicative of a low response or no response to treatment.

In accordance with this example, the disease may be selected from the group consisting of multiple sclerosis (MS), rheumatoid arthritis (RA), Type I diabetes (IDDM), scleroderma, Con A hepatitis, atopic dermatitis, asthma, allergic rhinitis and allergy.

Alternatively, another example of the present invention provides a process for accurately determining the likelihood that a subject will or will not respond to treatment of one or more bacterial or viral infections with an immunomodulatory composition, said process comprising performing a method according to any example hereof for accurately determining a likely response of a subject to treatment with an immunomodulatory composition, and determining a response for the subject selected from the group consisting of:

- (i) a response comprising enhanced clearance of a virus or bacterium or a reduction in virus titer or bacterial count or a change in other health characteristic of the subject related to reduced virus titer or bacterial count or enhanced clearance, wherein said response is indicative of a response to treatment; and
- (ii) a failure to clear of a virus or bacteria or to reduce virus titer or bacterial count or a change in a health characteristic of the subject related to said failure, wherein said response is indicative of a low response or no response to treatment.

In accordance with this example, the bacterium is a gram negative bacterium and/or the virus is a single-stranded RNA virus *e.g.*, a virus is selected from the group consisting of a human papillomavirus, a picornavirus, a flavivirus such as a hepatitis virus, an arenavirus, a togavirus, a bunyavirus, a filovirus, a paramyxovirus, a rhabdovirus, an

orthomyxovirus, and a coronavirus. Alternatively, the virus is a DNA virus *e.g.*, a herpesvirus.

Yet another example of the invention provides a process for accurately determining the  
5 likelihood that a subject will or will not respond to treatment of HCV infection with an immunomodulatory composition, said process comprising performing a method according to any example hereof for accurately determining a likely response of a subject to treatment with an immunomodulatory composition, and determining a response for the subject selected from the group consisting of:

- 10 (i) a response comprising enhanced clearance of HCV or a reduction in HCV titer or a change in other health characteristic of the subject related to reduced virus titer or enhanced clearance, wherein said response is indicative of a response to treatment; and  
(ii) a failure to clear HCV or to reduce HCV titer or a change in a health characteristic of the subject related to said failure, wherein said response is indicative of a low  
15 response or no response to treatment.

In yet another example, the present invention provides a process for accurately determining the likelihood that a subject will or will not respond to treatment of HCV infection with an immunomodulatory composition comprising an IFN or a derivative  
20 thereof and ribavirin or a derivative thereof, said process comprising performing a method according to any example hereof for accurately determining a likely response of a subject to treatment with an immunomodulatory composition, and determining a response for the subject selected from the group consisting of:

- (i) a response comprising enhanced clearance of HCV or a reduction in HCV titer or  
25 a change in other health characteristic of the subject related to reduced virus titer or enhanced clearance, wherein said response is indicative of a response to treatment; and  
(ii) a failure to clear HCV or to reduce HCV titer or a change in a health characteristic of the subject related to said failure, wherein said response is indicative of a low response or no response to treatment.

In these foregoing examples, the immunomodulatory composition may comprise one or more IFNs and/or one or more derivatives of said one or more of said IFNs as described according to any other example hereof. Alternatively, or in addition, the immunomodulatory composition may comprise one or more guanosine analogs and/or  
5 one or more derivatives of said one or more of said guanosine analogs according to any other example hereof.

In another example, the present invention provides a process for selecting a subject in need of treatment with an immunomodulatory composition, said process comprising:

- 10 (i) exposing a sample comprising cells obtained from the subject to the immunomodulatory composition *in vitro*; and
- (ii) performing a prognostic method or process as described according to any example hereof on the sample to thereby identify a subject likely to respond to treatment with the immunomodulatory composition; and
- 15 (iii) administering or recommending an immunomodulatory composition to a subject likely to respond to treatment.

In another example, the present invention provides a process for selecting a subject in need of treatment with an immunomodulatory composition, said process comprising:

- 20 (i) exposing a sample comprising cells obtained from the subject to the immunomodulatory composition *in vitro*; and
- (ii) performing a prognostic method or process as described according to any example hereof on the sample to thereby identify a subject likely to not respond to treatment with the immunomodulatory composition or likely to provide a low response to  
25 treatment; and
- (iii) administering or recommending an alternative therapy to the immunomodulatory composition.

This selection process is readily-performed on a sample from a subject that has not  
30 been previously administered with the immunomodulatory composition, or for determining whether or not to continue treatment in a subject who has received prior *in*

*vivo* administration of the immunomodulatory composition. This method is particularly well-suited to determining the effect of an immunomodulatory composition on a sample from a subject infected with HCV. In this example, the immunomodulatory composition may comprise one or more IFNs and/or one or more derivatives of said one or more of said IFNs as described according to any other example hereof. Alternatively, or in addition, the immunomodulatory composition may comprise one or more guanosine analogs and/or one or more derivatives of said one or more of said guanosine analogs according to any other example hereof. Exemplary samples comprise peripheral blood mononuclear cells.

10

In a related example, the present invention provides a process for treating an HCV-infected subject, comprising: (i) performing the *ex vivo* selection process on a sample from a subject; and (ii) either (a) administering or recommending a therapeutically effective amount of an immunomodulatory composition comprising an IFN to the subject or shortening the length of treatment with said immunomodulatory composition if the subject is likely to respond to treatment, or (b) administering or recommending an alternative therapy or extending the length of treatment with said immunomodulatory composition if the subject is not likely to respond to treatment or likely to produce a low response to treatment.

20

In another example, the present invention provides a process for determining a predisposition in a subject to a chronic HCV infection, said process comprising performing a prognostic method as described herein to thereby identify a subject likely to not respond to treatment with an immunomodulatory composition or likely to provide a low response to treatment, and determining that the subject has a predisposition to chronic HCV infection.

In yet another example, the present invention provides methods of treatment employing the prognostic test described herein. For example, the invention provides a process comprising: (i) performing a prognostic method or process as described according to any example hereof; and (ii) administering or recommending an immunomodulatory

30

composition to a subject. In another example, such a process comprises: (i) obtaining results of a prognostic method or process as described according to any example hereof; and (ii) administering or recommending an immunomodulatory composition to a subject.

5

A further example of the present invention provides a kit comprising a plurality of isolated nucleic acids and/or a plurality of antibodies and/or a plurality of peptides for performing a prognostic method or process according to any example hereof. In one example, the nucleic acids each comprise a fragment of an HLA-C1 allele and/or a  
10 fragment of a HLA-C2 allele, in combination with a fragment of a rs7248668 allele, and optionally one or more fragments of one or more other IL28B alleles listed in Table 1, wherein said fragments are capable of distinguishing between LR and HR alleles at the same locus *e.g.*, by virtue of comprising nucleotide sequences set forth herein or complementary thereto, or by virtue of being contained within said nucleotide  
15 sequences.

In another example, the kit comprises antibodies that bind selectively to HLA-C1 and/or antibodies that bind selectively to HLA-C2 in combination with antibodies that bind selectively to rs7248668-HR and/or antibodies that bind selectively to rs7248668-  
20 LR. Optionally, the kit further comprises one or more additional antibodies that bind selectively to an allelic variant of IL28B set forth in Table 1 and are capable of distinguishing between the other allelic variant at the same locus, wherein the allelic variant is other than an rs7248668 allele.

25 In another example, the kit comprises peptides that bind selectively to HLA-C1 and/or peptides that bind selectively to HLA-C2 in combination with peptides that bind selectively to rs7248668-HR and/or peptides that bind selectively to rs7248668-LR. Optionally, the kit further comprises one or more additional peptides that bind selectively to an allelic variant of IL28B set forth in Table 1 and are capable of  
30 distinguishing between the other allelic variant at the same locus, wherein the allelic variant is other than an rs7248668 allele.

The plurality of nucleic acids, peptides or antibodies may be arrayed *e.g.*, on a solid substrate. Preferably, the kit at least comprises a plurality of nucleic acids comprising sequences derived from HLA-C2 alleles in combination with nucleic acids comprising  
5 sequences derived from the rs7248668-LR allele, and at least one or two or three or four or five or six or seven or eight or nine or ten additional nucleic acids each comprising sequences derived from IL28B-LR alleles.

A further example provides for the use of a plurality of isolated nucleic acid or peptides  
10 or antibodies as described according to any example hereof in the manufacture of a kit or solid substrate for performing a prognostic method or process according to any example hereof.

### 3. *General*

15 Unless the context requires otherwise or specifically stated to the contrary, integers, steps, or elements of the invention recited herein as singular integers, steps or elements clearly encompass both singular and plural forms of the recited integers, steps or elements.

20 The designation of nucleotide residues referred to herein are those recommended by the IUPAC-IUB Biochemical Nomenclature Commission, wherein A represents Adenine, C represents Cytosine, G represents Guanine, T represents Thymine, Y represents a pyrimidine residue, R represents a purine residue, M represents Adenine or Cytosine, K represents Guanine or Thymine, S represents Guanine or Cytosine, W represents  
25 Adenine or Thymine, H represents a nucleotide other than Guanine, B represents a nucleotide other than Adenine, V represents a nucleotide other than Thymine, D represents a nucleotide other than Cytosine and N represents any nucleotide residue.

As used herein the term "derived from" shall be taken to indicate that a specified  
30 integer may be obtained from a particular source albeit not necessarily directly from that source.

Throughout this specification, unless the context requires otherwise, the word "comprise", or variations such as "comprises" or "comprising", will be understood to imply the inclusion of a stated step or element or integer or group of steps or elements  
5 or integers but not the exclusion of any other step or element or integer or group of elements or integers.

Throughout this specification, unless specifically stated otherwise or the context requires otherwise, reference to a single step, composition of matter, group of steps or  
10 group of compositions of matter shall be taken to encompass one and a plurality (*i.e.* one or more) of those steps, compositions of matter, groups of steps or group of compositions of matter.

Each embodiment described herein is to be applied *mutatis mutandis* to each and every  
15 other embodiment unless specifically stated otherwise.

Those skilled in the art will appreciate that the invention described herein is susceptible to variations and modifications other than those specifically described. It is to be understood that the invention includes all such variations and modifications. The  
20 invention also includes all of the steps, features, compositions and compounds referred to or indicated in this specification, individually or collectively, and any and all combinations or any two or more of said steps or features.

The present invention is not to be limited in scope by the specific embodiments  
25 described herein, which are intended for the purpose of exemplification only. Functionally-equivalent products, compositions and methods are clearly within the scope of the invention, as described herein.

The present invention is performed without undue experimentation using, unless  
30 otherwise indicated, conventional techniques of molecular biology, developmental biology, mammalian cell culture, recombinant DNA technology, histochemistry and

immunohistochemistry and immunology. Such procedures are described, for example, in the following texts that are incorporated by reference:

1. Sambrook, Fritsch & Maniatis, *Molecular Cloning: A Laboratory Manual*, Cold  
5 Spring Harbor Laboratories, New York, Second Edition (1989), whole of Vols I, II, and III;
2. *DNA Cloning: A Practical Approach*, Vols. I and II (D. N. Glover, ed., 1985), IRL Press, Oxford, whole of text;
3. *Oligonucleotide Synthesis: A Practical Approach* (M. J. Gait, ed., 1984) IRL  
10 Press, Oxford, whole of text, and particularly the papers therein by Gait, pp1-22; Atkinson *et al.*, pp35-81; Sproat *et al.*, pp 83-115; and Wu *et al.*, pp 135-151;
4. *Nucleic Acid Hybridization: A Practical Approach* (B. D. Hames & S. J. Higgins, eds., 1985) IRL Press, Oxford, whole of text;

15 Brief description of the drawings

Figure 1 provides a graphical representation showing combined expression of IFN- $\lambda$ 2 and IFN- $\lambda$ 3 (y-axis) as determined by RT-PCR for healthy controls having different genotypes at rs8099917 (x-axis). Data show that expression of IFN- $\lambda$ 2 and IFN- $\lambda$ 3 is reduced in patients that are homozygous for the low response (LR) G allele at this locus  
20 compared to those patients that are homozygous for the high response (HR) T allele at this locus, and for G/T heterozygotes. The data further suggest functional significance of the rs8099917 SNP in therapeutic response.

Figure 2 provides a graphical representation showing the percentage of individuals who  
25 have a low response (LR) or high response (HR) to antiviral therapy in a cohort of about 300 *e.g.*, 301, HCV-infected subjects. Subjects genotyped for IL28B alleles alone were either homozygous for a HR allele linked to IL28B at rs8099917 (TT on the x-axis) or had at least one LR allele linked to IL28B at rs8099917 (G\* on the x-axis). Subjects genotyped for HLA-C alleles alone were either homozygous for HLA-C2  
30 (C2C2 on the x-axis) or had at least one HLA-C1 allele (not C2C2 on the x-axis). Subjects genotyped for HLA-C alleles and at least one IL28B allele were either

homozygous for HLA-C2 and at least one LR allele linked to IL28B at rs8099917 (C2C2G\* on the x-axis), or they had at least one HLA-C1 allele and no LR allele linked to IL28B at rs8099917 (Not C2C2G\* on the x-axis). Data indicate that HLA-C2 homozygous individuals respond poorly to therapy, and that its effect is more marked if  
5 the subjects are also carrier of the LR allele linked to IL28B at rs8099917 ( $p < 0.0001$ ).

Figure 3 provides a graphical representation showing the percentage non-SVR rates of the three HLA-C/rs8099917 genotype combinations in stage 1, replication stage and combined cohort as described in Example 4.

10

Figure 4 provides a graphical representation showing the percentage non-SVR rates of HLA-C, IL28B and combined HLA-C/IL28B genotypes in the combined cohort as described in Example 4.

15 Figure 5 provides a graphical representation of the SNP Selection Scheme employed in this study

Figure 6 provides a graphical representation of the results of allele-based association tests at the locations of variants as called by CRISP using pooled MPS data.

20

Figure 7 provides a graphical representation of odds ratios for the eighteen individually genotyped SNPs under four different genetic models.

Figure 8 provides a graphical representation of location and D' values for the SNPs  
25 genotyped in this study. Linkage disequilibrium blocks determined from the cohort data using Haploview. HapMap SNPs genotyped in multiple populations are shown in the header map in each case.

Figure 9 provides a graphical representation of the putative Transcription Factor and  
30 Methylation Sites on IL28B Haplotypes. The 6 haplotypes identified using Haploview are shown. SNPs changing CpG sites in the region identified as methylated are boxed

in red. *In silico* predictions of different haplotype transcription factor binding sites were identified using Ali Baba, and serve as a proof of principle that the haplotype sequence differences are sufficient to alter response to transcription factors.

5 Detailed description of the preferred embodiments

*Markers associated with a disease or disorder*

In one example, a marker is at least one HLA-C allele *e.g.*, at least one HLA-C1 allele or at least one HLA-C2 allele. In another example, a marker of the present invention is  
10 at least one IL28B allele *e.g.*, at least one HR allele linked to IL28B or at least one LR allele linked to IL28B or a IL28B haplotype block.

Protein markers for HLA-C1 alleles or HLA-C2 alleles are described by Alper *et al.*, *J. Exp. Med.* 144, p1111 *et seq.* (1976) or Hobart *et al.*, *J. Immunol.* 116, p1736 *et seq.*  
15 (1976) incorporated herein by reference.

Alternatively or in addition, markers comprising or consisting of nucleic acid are employed *e.g.*, DNA or RNA or a combination of RNA or DNA or modified RNA comprising one or more non-nucleic acid components or modified DNA comprising  
20 one or more non-nucleic acid components.

For example, HLA-C markers may comprise amplified nucleic acid comprising a polymorphism such as a SNP, microsatellite or INDEL. Standard means are employed to determine HLA-C alleles. In one example, the Class I SSP ARMS-PCR method is  
25 employed essentially as described by Tonks *et al.*, *Tissue Antigens* 53, 175–183, 1999 which is incorporated herein by reference. In another example, HLA-C alleles are amplified using amplification primers comprising one or more hypervariable regions of HLA-C1 or HLA-C2 genes, such as the hypervariable region of exon 2 and/or exon 3. For example, primers specific for exon 2, codon 45 and/or exon 3, codon 182 may be  
30 employed. Alternatively, or in addition, amplification primers specific for nucleic acid encoding the HLA-Cw  $\alpha$ 1 domain *e.g.*, comprising polymorphisms for codon 77 and/or

codon 80 are employed to distinguish HLA-C1 from HLA-C2. Specificity of sequence-specific oligonucleotide probes (SSOPs) is confirmed by using B lymphoblastoid cell lines as positive controls for the HLA-C alleles.

- 5 Specific microsatellite markers for HLA-C alleles are described *e.g.*, by Tamiya *et al.*, *Tissue Antigens* 51, 337-346, (1998); Tamiya *et al.*, *Tissue Antigens* 52, (1999); Jenisch *et al.*, *Am. J. Hum. Genet.* 63, 191-199 (1998); Marshall *et al.*, *Hum. Immunol.* 38, 24-29 (1993); or Krishnan *et al.*, *Genomics* 30, 53-58 (1995), each of which is incorporated by reference.

10

Exemplary SNP markers for HLA-C alleles are from the region containing HLA-B and HLA-C *i.e.*, the so-called beta-block as described *e.g.*, by Gaudieri *et al.*, *Genome Res.* 10, 1579-1586 (2000) and the references cited therein, each of which is incorporated herein by reference.

15

- Preferred IL28B SNP markers will comprise a sequence set forth in the Sequence Listing or complementary thereto. Such a nucleic acid marker comprises, for example, a polymorphism, an insertion into an IL28B gene or transcript thereof, a deletion from an IL28B gene or transcript thereof, a transcript of an IL28B gene or a fragment thereof  
20 or an alternatively spliced transcript of an IL28B or a fragment thereof, and includes copy number variants or inversions. The nucleotide substitution or deletion or insertion may be in the 5'-end of a gene, the 3'-end of a gene, in an exon of a gene or an intron of a gene. Alternatively, the nucleotide substitution or deletion or insertion may be in an intergenic region *i.e.*, between genes. A nucleotide substitution or deletion or  
25 insertion may modify gene expression and, without being bound by any theory or mode of action this modified expression may be associated with the development of a therapeutic response, or a non-response or low response.

- In one example, the method of the invention comprises detecting or determining the  
30 presence of a plurality of markers associated with a therapeutic response.

*Assay detection methods**(i) Nucleic acid marker detection*

As will be apparent to the skilled artisan a probe or primer capable of specifically detecting a marker that is associated with or causative of a therapeutic response, is any  
5 probe or primer that is capable of specifically hybridizing to the region of the genome that comprises said marker, or an expression product thereof. Accordingly, a nucleic acid marker is preferably at least about 8 nucleotides in length (for example, for detection using a locked nucleic acid (LNA) probe). To provide more specific hybridization, a marker is preferably at least about 15 nucleotides in length or more  
10 preferably at least 20 to 30 nucleotides in length. Such markers are particularly amenable to detection by nucleic acid hybridization-based detection means assays, such as, for example any known format of PCR or ligase chain reaction.

Generally, a method for detecting a nucleic acid marker comprises hybridizing an  
15 oligonucleotide to the marker linked to nucleic acid in a sample from a subject under moderate to high stringency conditions and detecting hybridization of the oligonucleotide using a detection means, such as for example, an amplification reaction or a hybridization reaction.

20 For the purposes of defining the level of stringency to be used in these diagnostic assays, a low stringency is defined herein as being a hybridization and/or a wash carried out in 6 x SSC buffer, 0.1% (w/v) SDS at 28°C, or equivalent conditions. A moderate stringency is defined herein as being a hybridization and/or washing carried out in 2 x SSC buffer, 0.1% (w/v) SDS at a temperature in the range 45°C to 65°C, or  
25 equivalent conditions. A high stringency is defined herein as being a hybridization and/or wash carried out in 0.1 x SSC buffer, 0.1% (w/v) SDS, or lower salt concentration, and at a temperature of at least 65°C, or equivalent conditions. Reference herein to a particular level of stringency encompasses equivalent conditions using wash/hybridization solutions other than SSC known to those skilled in the art.

Generally, the stringency is increased by reducing the concentration of SSC buffer, and/or increasing the concentration of SDS and/or increasing the temperature of the hybridization and/or wash. Those skilled in the art will be aware that the conditions for hybridization and/or wash may vary depending upon the nature of the hybridization matrix used to support the sample DNA, and/or the type of hybridization probe used.

In another example, stringency is determined based upon the temperature at which a probe or primer dissociates from a target sequence (*i.e.*, the probe or primers melting temperature or  $T_m$ ). Such a temperature may be determined using, for example, an equation or by empirical means. Several methods for the determination of the  $T_m$  of a nucleic acid are known in the art. For example the Wallace Rule determines the G + C and the T + A concentrations in the oligonucleotide and uses this information to calculate a theoretical  $T_m$  (Wallace *et al.*, *Nucleic Acids Res.* 6, 3543, 1979). Alternative methods, such as, for example, the nearest neighbour method are known in the art, and described, for example, in Howley, *et al.*, *J. Biol. Chem.* 254, 4876, Santa Lucia, *Proc. Natl. Acad. Sci. USA*, 95: 1460-1465, 1995 or Bresslauer *et al.*, *Proc. Natl. Acad. Sci. USA*, 83: 3746-3750, 1986. A temperature that is similar to (*e.g.*, within 5°C or within 10°C) or equal to the proposed denaturing temperature of a probe or primer is considered to be high stringency. Medium stringency is to be considered to be within 10°C to 20°C or 10°C to 15°C of the calculated  $T_m$  of the probe or primer.

a) Probe/primer design and production

As will be apparent to the skilled artisan, the specific probe or primer used in an assay of the present invention will depend upon the assay format used. Clearly, a probe or primer that is capable of preferentially or specifically hybridizing or annealing to or  
5 detecting the marker of interest is preferred. Methods for designing probes and/or primers for, for example, PCR or hybridization are known in the art and described, for example, in Dieffenbach and Dveksler (Eds) (In: PCR Primer: A Laboratory Manual, Cold Spring Harbor Laboratories, NY, 1995). Furthermore, several software packages are publicly available that design optimal probes and/or primers for a variety of assays,  
10 e.g. Primer 3 available from the Center for Genome Research, Cambridge, MA, USA. Probes and/or primers useful for detection of a marker associated with a therapeutic response, are assessed to determine those that do not form hairpins, self-prime or form primer dimers (e.g. with another probe or primer used in a detection assay).

15 Furthermore, a probe or primer (or the sequence thereof) is assessed to determine the temperature at which it denatures from a target nucleic acid (i.e. the melting temperature of the probe or primer, or  $T_m$ ). Methods of determining  $T_m$  are known in the art and described, for example, in Santa Lucia, *Proc. Natl. Acad. Sci. USA*, 95: 1460-1465, 1995 or Bresslaue *et al.*, *Proc. Natl. Acad. Sci. USA*, 83: 3746-3750, 1986.

20

A primer or probe useful for detecting a SNP or mutation in an allele specific PCR assay or a ligase chain reaction assay is designed such that the 3' terminal nucleotide hybridizes to the site of the SNP or mutation. The 3' terminal nucleotide may be any of the nucleotides known to be present at the site of the SNP or mutation. When  
25 complementary nucleotides occur in the probe or primer and at the site of the polymorphism the 3' end of the probe or primer hybridizes completely to the marker of interest and facilitates amplification, for example, PCR amplification or ligation to another nucleic acid. Accordingly, a probe or primer that completely hybridizes to the target nucleic acid produces a positive result in an assay.

30

In another example, a primer useful for a primer extension reaction is designed such that it preferentially or specifically hybridizes to a region adjacent to a specific nucleotide of interest, *e.g.* a SNP or mutation.

- 5 Whilst the specific hybridization of a probe or primer may be estimated by determining the degree of homology of the probe or primer to any nucleic acid using software, such as, for example, BLAST, the specificity of a probe or primer can only be determined empirically using methods known in the art.
- 10 A locked nucleic acid (LNA) or protein-nucleic acid (PNA) probe or a molecular beacon useful, for example, for detection of a SNP or mutation or microsatellite by hybridization is at least about 8 to 12 nucleotides in length. Preferably, the nucleic acid, or derivative thereof, that hybridizes to the site of the SNP or mutation or microsatellite is positioned at approximately the centre of the probe, thereby facilitating selective
- 15 hybridization and accurate detection.

Methods for producing/synthesizing a probe or primer of the present invention are known in the art. For example, oligonucleotide synthesis is described, in Gait (Ed) (*In: Oligonucleotide Synthesis: A Practical Approach*, IRL Press, Oxford, 1984). For

20 example, a probe or primer may be obtained by biological synthesis (*eg.* by digestion of a nucleic acid with a restriction endonuclease) or by chemical synthesis. For short sequences (up to about 100 nucleotides) chemical synthesis is preferable.

For longer sequences standard replication methods employed in molecular biology are

25 useful, such as, for example, the use of M13 for single stranded DNA as described by J. Messing (1983) *Methods Enzymol*, 101, 20-78.

Other methods for oligonucleotide synthesis include, for example, phosphotriester and phosphodiester methods (Narang, *et al. Meth. Enzymol* 68: 90, 1979) and synthesis on a

30 support (Beaucage, *et al Tetrahedron Letters* 22: 1859-1862, 1981) as well as phosphoramidate technique, Caruthers, M. H., *et al.*, "Methods in Enzymology," Vol.

- 154, pp. 287-314 (1988), and others described in "Synthesis and Applications of DNA and RNA," S. A. Narang, editor, Academic Press, New York, 1987, and the references contained therein.
- 5 LNA synthesis is described, for example, in Nielsen *et al*, *J. Chem. Soc. Perkin Trans., I*: 3423, 1997; Singh and Wengel, *Chem. Commun.* 1247, 1998. While, PNA synthesis is described, for example, in Egholm *et al.*, *Am. Chem. Soc.*, 114: 1895, 1992; Egholm *et al.*, *Nature*, 365: 566, 1993; and Orum *et al.*, *Nucl. Acids Res.*, 21: 5332, 1993.
- 10 In one example, the probe or primer comprises one or more detectable markers. For example, the probe or primer comprises a fluorescent label such as, for example, fluorescein (FITC), 5,6-carboxymethyl fluorescein, Texas red, nitrobenz-2-oxa-1,3-diazol-4-yl (NBD), coumarin, dansyl chloride, rhodamine, 4'-6-diamidino-2-phenylindole (DAPI), and the cyanine dyes Cy3, Cy3.5, Cy5, Cy5.5 and Cy7,
- 15 fluorescein (5-carboxyfluorescein-N-hydroxysuccinimide ester), rhodamine (5,6-tetramethyl rhodamine). The absorption and emission maxima, respectively, for these fluors are: FITC (490 nm; 520 nm), Cy3 (554 nm; 568 nm), Cy3.5 (581 nm; 588 nm), Cy5 (652 nm; 672 nm), Cy5.5 (682 nm; 703 nm) and Cy7 (755 nm; 778 nm).
- 20 Alternatively, the probe or primer is labeled with, for example, a fluorescent semiconductor nanocrystal (as described, for example, in US 6,306,610), a radiolabel or an enzyme (*e.g.* horseradish peroxidase (HRP), alkaline phosphatase (AP) or  $\beta$ -galactosidase).
- 25 Such detectable labels facilitate the detection of a probe or primer, for example, the hybridization of the probe or primer or an amplification product produced using the probe or primer. Methods for producing such a labeled probe or primer are known in the art. Furthermore, commercial sources for the production of a labeled probe or primer will be known to the skilled artisan, *e.g.*, Sigma-Genosys, Sydney, Australia.

The present invention additionally contemplates the use a probe or primer as described herein in the manufacture of a diagnostic reagent for diagnosing or determining a predisposition to a therapeutic response.

5    b) Detection methods

Methods for detecting nucleic acids are known in the art and include for example, hybridization based assays, amplification based assays and restriction endonuclease based assays. For example, a change in the sequence of a region of the genome or an expression product thereof, such as, for example, an insertion, a deletion, a  
10   transversion, a transition, alternative splicing or a change in the preference of or occurrence of a splice form of a gene is detected using a method, such as, polymerase chain reaction (PCR) strand displacement amplification, ligase chain reaction, cycling probe technology or a DNA microarray chip amongst others.

15   Methods of PCR are known in the art and described, for example, in Dieffenbach (Ed) and Dveksler (Ed) (*In: PCR Primer: A Laboratory Manual*, Cold Spring Harbor Laboratories, NY, 1995). Generally, for PCR two non-complementary nucleic acid primer molecules comprising at least about 20 nucleotides in length, and more preferably at least 30 nucleotides in length are hybridized to different strands of a  
20   nucleic acid template molecule, and specific nucleic acid molecule copies of the template are amplified enzymatically. PCR products may be detected using electrophoresis and detection with a detectable marker that binds nucleic acids. Alternatively, one or more of the oligonucleotides are labeled with a detectable marker (e.g. a fluorophore) and the amplification product detected using, for example, a  
25   lightcycler (Perkin Elmer, Wellesley, MA, USA). Clearly, the present invention also encompasses quantitative forms of PCR, such as, for example, Taqman assays.

Strand displacement amplification (SDA) utilizes oligonucleotides, a DNA polymerase and a restriction endonuclease to amplify a target sequence. The oligonucleotides are  
30   hybridized to a target nucleic acid and the polymerase used to produce a copy of this region. The duplexes of copied nucleic acid and target nucleic acid are then nicked with

an endonuclease that specifically recognizes a sequence at the beginning of the copied nucleic acid. The DNA polymerase recognizes the nicked DNA and produces another copy of the target region at the same time displacing the previously generated nucleic acid. The advantage of SDA is that it occurs in an isothermal format, thereby  
5 facilitating high-throughput automated analysis.

Ligase chain reaction (described in EU 320,308 and US 4,883,750) uses at least two oligonucleotides that bind to a target nucleic acid in such a way that they are adjacent. A ligase enzyme is then used to link the oligonucleotides. Using thermocycling the  
10 ligated oligonucleotides then become a target for further oligonucleotides. The ligated fragments are then detected, for example, using electrophoresis, or MALDI-TOF. Alternatively, or in addition, one or more of the probes is labeled with a detectable marker, thereby facilitating rapid detection.

15 Cycling Probe Technology uses chimeric synthetic probe that comprises DNA-RNA-DNA that is capable of hybridizing to a target sequence. Upon hybridization to a target sequence the RNA-DNA duplex formed is a target for RNase H thereby cleaving the probe. The cleaved probe is then detected using, for example, electrophoresis or MALDI-TOF.

20

In a preferred example, a marker that is associated with or causative of a therapeutic response, occurs within a protein coding region of a genomic gene (*e.g.* an IFN- $\Lambda$ 3 gene) and is detectable in mRNA encoded by that gene. For example, such a marker may be an alternate splice-form of a mRNA encoded by a genomic gene (*e.g.* a splice  
25 form not observed in a normal and/or healthy subject, or, alternatively, an increase or decrease in the level of a splice form in a subject that carries the marker). Such a marker may be detected using, for example, reverse-transcriptase PCR (RT-PCR), transcription mediated amplification (TMA) or nucleic acid sequence based amplification (NASBA), although any mRNA or cDNA based hybridization and/or  
30 amplification protocol is clearly amenable to the instant invention.

Methods of RT-PCR are known in the art and described, for example, in Dieffenbach (Ed) and Dveksler (Ed) (*In: PCR Primer: A Laboratory Manual*, Cold Spring Harbor Laboratories, NY, 1995).

- 5 Methods of TMA or self-sustained sequence replication (3SR) use two or more oligonucleotides that flank a target sequence, a RNA polymerase, RNase H and a reverse transcriptase. One oligonucleotide (that also comprises a RNA polymerase binding site) hybridizes to an RNA molecule that comprises the target sequence and the reverse transcriptase produces cDNA copy of this region. RNase H is used to digest the  
10 RNA in the RNA-DNA complex, and the second oligonucleotide used to produce a copy of the cDNA. The RNA polymerase is then used to produce a RNA copy of the cDNA, and the process repeated.

- NASBA systems rely on the simultaneous activity of three enzymes (a reverse  
15 transcriptase, RNase H and RNA polymerase) to selectively amplify target mRNA sequences. The mRNA template is transcribed to cDNA by reverse transcription using an oligonucleotide that hybridizes to the target sequence and comprises a RNA polymerase binding site at its 5' end. The template RNA is digested with RNase H and double stranded DNA is synthesized. The RNA polymerase then produces multiple  
20 RNA copies of the cDNA and the process is repeated.

- Clearly, the hybridization to and/or amplification of a marker associated with a therapeutic response, using any of these methods is detectable using, for example, electrophoresis and/or mass spectrometry. In this regard, one or more of the  
25 probes/primers and/or one or more of the nucleotides used in an amplification reactions may be labeled with a detectable marker to facilitate rapid detection of a marker, for example, marker as described *supra*, e.g., a fluorescent label (e.g. Cy5 or Cy3) or a radioisotope (e.g.  $^{32}\text{P}$ ).

- 30 Alternatively, amplification of a nucleic acid may be continuously monitored using a melting curve analysis method, such as that described in, for example, US 6,174,670.

In a one exemplified form of the invention, a marker associated with a therapeutic response, comprises a single nucleotide change. Methods of detecting single nucleotide changes are known in the art, and reviewed, for example, in Landegren *et al*, *Genome Research* 8: 769-776, 1998.

For example, a single nucleotide changes that introduces or alters a sequence that is a recognition sequence for a restriction endonuclease is detected by digesting DNA with the endonuclease and detecting the fragment of interest using, for example, Southern blotting (described in Ausubel *et al* (In: Current Protocols in Molecular Biology. Wiley Interscience, ISBN 047 150338, 1987) and Sambrook *et al* (In: Molecular Cloning: Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratories, New York, Third Edition 2001)). Alternatively, a nucleic acid amplification method described *supra*, is used to amplify the region surrounding the single nucleotide changes. The amplification product is then incubated with the endonuclease and any resulting fragments detected, for example, by electrophoresis, MALDI-TOF or PCR.

The direct analysis of the sequence of polymorphisms of the present invention can be accomplished using either the dideoxy chain termination method or the Maxam-Gilbert method (see Sambrook *et al.*, Molecular Cloning, A Laboratory Manual (2nd Ed., CSHP, New York 1989); Zyskind *et al.*, Recombinant DNA Laboratory Manual, (Acad. Press, 1988)).

Alternatively, a single nucleotide change is detected using single stranded conformational polymorphism (SSCP) analysis. SSCP analysis relies upon the formation of secondary structures in nucleic acids and the sequence dependent nature of these secondary structures. In one form of this analysis an amplification method, such as, for example, a method described *supra*, is used to amplify a nucleic acid that comprises a single nucleotide change. The amplified nucleic acids are then denatured, cooled and analyzed using, for example, non-denaturing polyacrylamide gel electrophoresis, mass spectrometry, or liquid chromatography (*e.g.* HPLC or dHPLC).

Regions that comprise different sequences form different secondary structures, and as a consequence migrate at different rates through, for example, a gel and/or a charged field. Clearly, a detectable marker may be incorporated into a probe/primer useful in SSCP analysis to facilitate rapid marker detection.

5

Alternatively, any nucleotide changes are detected using, for example, mass spectrometry or capillary electrophoresis. For example, amplified products of a region of DNA comprising a single nucleotide change from a test sample are mixed with amplified products from a normal/healthy individual. The products are denatured and  
10 allowed to re-anneal. Clearly those samples that comprise a different nucleotide at the position of the single nucleotide change will not completely anneal to a nucleic acid molecule from a normal/healthy individual thereby changing the charge and/or conformation of the nucleic acid, when compared to a completely annealed nucleic acid. Such incorrect base pairing is detectable using, for example, mass spectrometry.

15

Mass spectrometry is also useful for detecting the molecular weight of a short amplified product, wherein a nucleotide change causes a change in molecular weight of the nucleic acid molecule (such a method is described, for example, in US 6,574,700).

20 Allele specific PCR (as described, for example, In Liu *et al*, *Genome Research*, 7: 389-398, 1997) is also useful for determining the presence of one or other allele of a single nucleotide change. An oligonucleotide is designed, in which the most 3' base of the oligonucleotide hybridizes with the single nucleotide change. During a PCR reaction, if the 3' end of the oligonucleotide does not hybridize to a target sequence, little or no  
25 PCR product is produced, indicating that a base other than that present in the oligonucleotide is present at the site of single nucleotide change in the sample. PCR products are then detected using, for example, gel or capillary electrophoresis or mass spectrometry.

30 Primer extension methods (described, for example, in Dieffenbach (Ed) and Dveksler (Ed) (*In*: PCR Primer: A Laboratory Manual, Cold Spring Harbor Laboratories, NY,

1995)) are also useful for the detection of a single nucleotide change. An oligonucleotide that hybridizes to the region of a nucleic acid adjacent to the single nucleotide change. This oligonucleotide is then used in a primer extension protocol with a polymerase and a free nucleotide diphosphate that corresponds to either or any  
5 of the possible bases that occur at the single nucleotide change. Preferably the nucleotide-diphosphate is labeled with a detectable marker (e.g. a fluorophore). Following primer extension, unbound labeled nucleotide diphosphates are removed, e.g. using size exclusion chromatography or electrophoresis, or hydrolyzed, using for example, alkaline phosphatase, and the incorporation of the labeled nucleotide into the  
10 oligonucleotide is detected, indicating the base that is present at the site of the single nucleotide change. Alternatively, or in addition, as exemplified herein primer extension products are detected using mass spectrometry (e.g. MALDI-TOF).

Clearly, the present invention extends to high-throughput forms primer extension  
15 analysis, such as, for example, minisequencing (Sy Vämen *et al.*, *Genomics* 9: 341-342, 1995). In such a method, a probe or primer (or multiple probes or primers) are immobilized on a solid support (e.g. a glass slide). A biological sample comprising nucleic acid is then brought into direct contact with the probe/s or primer/s, and a primer extension protocol performed with each of the free nucleotide bases labeled  
20 with a different detectable marker. The nucleotide present at a single nucleotide change or a number of single nucleotide changes is then determined by determining the detectable marker bound to each probe and/or primer.

Fluorescently labeled locked nucleic acid (LNA) molecules or fluorescently labeled  
25 protein-nucleic acid (PNA) molecules are useful for the detection of SNPs (as described in Simeonov and Nikiforov, *Nucleic Acids Research*, 30(17): 1-5, 2002). LNA and PNA molecules bind, with high affinity, to nucleic acid, in particular, DNA. Fluorophores (in particular, rhodamine or hexachlorofluorescein) conjugated to the LNA or PNA probe fluoresce at a significantly greater level upon hybridization of the  
30 probe to target nucleic acid. However, the level of increase of fluorescence is not enhanced to the same level when even a single nucleotide mismatch occurs.

Accordingly, the degree of fluorescence detected in a sample is indicative of the presence of a mismatch between the LNA or PNA probe and the target nucleic acid, such as, in the presence of a SNP. Preferably, fluorescently labeled LNA or PNA technology is used to detect a single base change in a nucleic acid that has been  
5 previously amplified using, for example, an amplification method described *supra*.

As will be apparent to the skilled artisan, LNA or PNA detection technology is amenable to a high-throughput detection of one or more markers immobilizing an LNA or PNA probe to a solid support, as described in Orum *et al.*, *Clin. Chem.* 45: 1898-  
10 1905, 1999.

Similarly, Molecular Beacons are useful for detecting single nucleotide changes directly in a sample or in an amplified product (see, for example, Mhlang and Malmberg, *Methods* 25: 463-471, 2001). Molecular beacons are single stranded nucleic  
15 acid molecules with a stem-and-loop structure. The loop structure is complementary to the region surrounding the single nucleotide change of interest. The stem structure is formed by annealing two "arms," complementary to each other, that are on either side of the probe (loop). A fluorescent moiety is bound to one arm and a quenching moiety to the other arm that suppresses any detectable fluorescence when the molecular beacon  
20 is not bound to a target sequence. Upon binding of the loop region to its target nucleic acid the arms are separated and fluorescence is detectable. However, even a single base mismatch significantly alters the level of fluorescence detected in a sample. Accordingly, the presence or absence of a particular base at the site of a single nucleotide change is determined by the level of fluorescence detected.

25

A single nucleotide change can also be identified by hybridization to nucleic acid arrays, an example of which is described in WO 95/11995. WO 95/11995 also describes subarrays that are optimized for detection of a variant form of a pre-characterized polymorphism. Such a sub-array contains probes designed to be  
30 complementary to a second reference sequence, which is an allelic variant of the first reference sequence. The second group of probes is designed by the same principles,

except that the probes exhibit complementarity to the second reference sequence. The inclusion of a second group (or further groups) can be particularly useful for analyzing short sub-sequences of the primary reference sequence in which multiple mutations are expected to occur within a short distance commensurate with the length of the probes  
5 (e.g., two or more mutations within 9 to 21 bases).

Clearly the present invention encompasses other methods of detecting a single nucleotide change that is associated with a therapeutic response, such as, for example, SNP microarrays (available from Affymetrix, or described, for example, in US  
10 6,468,743 or Hacia *et al*, *Nature Genetics*, 14: 441, 1996), Taqman assays (as described in Livak *et al*, *Nature Genetics*, 9: 341-342, 1995), solid phase minisequencing (as described in Syvämen *et al*, *Genomics*, 13: 1008-1017, 1992), minisequencing with FRET (as described in Chen and Kwok, *Nucleic Acids Res.* 25: 347-353, 1997) or pyrominisequencing (as reviewed in Landegren *et al*, *Genome Res.*, 8(8): 769-776,  
15 1998).

In a preferred example, a single nucleotide change associated with a therapeutic response, is detected using a Taqman assay essentially as described by Corder *et al*, *Science*, 261: 921-923.

20

(ii) *Protein marker detection*

a) Serological markers

Serological markers may be employed conveniently to determine HLA-C serotypes arising by expression of HLA-C1 alleles or HLA-C2 alleles e.g., by isoelectric  
25 focussing (IEF) of sera and C2-specific hemolytic agarose overlay as described by Alper *et al*, *J. Exp. Med.* 144, p1111 *et seq.* (1976) or Hobart *et al*, *J. Immunol.* 116, p1736 *et seq.* (1976) incorporated herein by reference.

b) Antibodies

30 Methods for detecting polypeptides generally make use of a ligand or antibody that preferentially or specifically binds to the target polypeptide. As used herein the term

“ligand” shall be taken in its broadest context to include any chemical compound, polynucleotide, peptide, protein, lipid, carbohydrate, small molecule, natural product, polymer, etc. that is capable of selectively binding, whether covalently or not, to one or more specific sites on a polypeptide encoded by a gene linked to a SNP of Table 1.

- 5 The ligand may bind to its target via any means including hydrophobic interactions, hydrogen bonding, electrostatic interactions, van der Waals interactions, pi stacking, covalent bonding, or magnetic interactions amongst others. It is particularly preferred that a ligand is able to specifically bind to a specific form of a polypeptide marker.
- 10 As used herein, the term “antibody” refers to intact monoclonal or polyclonal antibodies, immunoglobulin (IgA, IgD, IgG, IgM, IgE) fractions, humanized antibodies, or recombinant single chain antibodies, as well as fragments thereof, such as, for example Fab, F(ab)2, and Fv fragments.
- 15 Antibodies are prepared by any of a variety of techniques known to those of ordinary skill in the art, and described, for example in, Harlow and Lane (*In: Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory, 1988). In one such technique, an immunogen comprising the antigenic polypeptide is initially injected into any one of a wide variety of animals (*e.g.*, mice, rats, rabbits, sheep, humans, dogs, pigs, chickens  
20 and goats). The immunogen is derived from a natural source, produced by recombinant expression means, or artificially generated, such as by chemical synthesis (*e.g.*, BOC chemistry or Fmoc chemistry). A peptide comprising any variant amino acid listed in Table 1 may be employed as an antigen for antibody production.
- 25 A peptide, polypeptide or protein is joined to a carrier protein, such as bovine serum albumin or keyhole limpet hemocyanin. The immunogen and optionally a carrier for the protein is injected into the animal host, preferably according to a predetermined schedule incorporating one or more booster immunizations, and blood collected from said the animals periodically. Optionally, the immunogen is injected in the presence of  
30 an adjuvant, such as, for example Freund’s complete or incomplete adjuvant, lysolecithin and dinitrophenol to enhance the subject’s immune response to the

immunogen. Monoclonal or polyclonal antibodies specific for the polypeptide are then purified from blood isolated from an animal by, for example, affinity chromatography using the polypeptide coupled to a suitable solid support.

- 5 Monoclonal antibodies specific for the antigenic polypeptide of interest are prepared, for example, using the technique of Kohler and Milstein, *Eur. J. Immunol.* 6:511-519, 1976, and improvements thereto. Briefly, these methods involve the preparation of immortal cell lines capable of producing antibodies having the desired specificity (*i.e.*, reactivity with the polypeptide of interest). Such cell lines are produced, for example,
- 10 from spleen cells obtained from an animal immunized as described *supra*. The spleen cells are immortalized by, for example, fusion with a myeloma cell fusion partner, preferably one that is syngeneic with the immunized animal. A variety of fusion techniques are known in the art, for example, the spleen cells and myeloma cells are combined with a nonionic detergent or electrofused and then grown in a selective
- 15 medium that supports the growth of hybrid cells, but not myeloma cells. A preferred selection technique uses HAT (hypoxanthine, aminopterin, and thymine) selection. After a sufficient time, usually about 1 to 2 weeks, colonies of hybrids are observed. Single colonies are selected and growth media in which the cells have been grown is tested for the presence of an antibody having binding activity against the polypeptide
- 20 (immunogen). Hybridomas having high reactivity and specificity are preferred.

Monoclonal antibodies are isolated from the supernatants of growing hybridoma colonies using methods such as, for example, affinity purification as described *supra*.

- 25 Various techniques are also known for enhancing antibody yield, such as injection of the hybridoma cell line into the peritoneal cavity of a suitable vertebrate host, such as a mouse. Monoclonal antibodies are then harvested from the ascites fluid or the blood of such an animal subject. Contaminants are removed from the antibodies by conventional techniques, such as chromatography, gel filtration, precipitation, and/or extraction. The
- 30 marker associated with neurodegeneration of this invention may be used in the purification process in, for example, an affinity chromatography step.

It is preferable that an immunogen used in the production of an antibody is one which is sufficiently antigenic to stimulate the production of antibodies that will bind to the immunogen and is preferably, a high titer antibody. In one example, an immunogen is an entire protein. In another example, an immunogen consists of a peptide representing a fragment of a polypeptide. Preferably an antibody raised to such an immunogen also recognizes the full-length protein from which the immunogen was derived, such as, for example, in its native state or having native conformation.

Alternatively, or in addition, an antibody raised against a peptide immunogen recognizes the full-length protein from which the immunogen was derived when the protein is denatured. By "denatured" is meant that conformational epitopes of the protein are disrupted under conditions that retain linear B cell epitopes of the protein. As will be known to a skilled artisan linear epitopes and conformational epitopes may overlap.

Alternatively, a monoclonal antibody is produced using a method such as, for example, a human B-cell hybridoma technique (Kozbar *et al.*, *Immunol. Today* 4:72, 1983), a EBV-hybridoma technique to produce human monoclonal antibodies (Cole *et al.* *Monoclonal Antibodies in Cancer Therapy*, 1985 Allen R. Bliss, Inc., pages 77-96), or screening of combinatorial antibody libraries (Huse *et al.*, *Science* 246:1275, 1989).

Such an antibody is then particularly useful in detecting the presence of a marker of a therapeutic response.

The methods described *supra* are also suitable for production of an antibody or antibody binding fragment as described herein according to any example.

#### c) Detection methods

In one example, the method of the invention detects the presence of a marker in a polypeptide, said marker being associated or causative of with a therapeutic response.

An amount, level or presence of a polypeptide is determined using any of a variety of techniques known to the skilled artisan such as, for example, a technique selected from the group consisting of, immunohistochemistry, immunofluorescence, an immunoblot, a Western blot, a dot blot, an enzyme linked immunosorbent assay (ELISA), radioimmunoassay (RIA), enzyme immunoassay, isoelectric focussing (IEF), fluorescence resonance energy transfer (FRET), matrix-assisted laser desorption/ionization time of flight (MALDI-TOF), electrospray ionization (ESI), mass spectrometry (including tandem mass spectrometry, *e.g.* LC MS/MS), biosensor technology, evanescent fiber-optics technology or protein chip technology.

In one example, an assay used to determine the amount or level of a protein is a semi-quantitative assay. In another example, an assay used to determine the amount or level of a protein in a quantitative assay.

15

Preferably, an amount of antibody or ligand bound to a marker of a therapeutic response is determined using an immunoassay. Preferably, using an assay selected from the group consisting of immunohistochemistry, immunofluorescence, enzyme linked immunosorbent assay (ELISA), fluorescence linked immunosorbent assay (FLISA) Western blotting, RIA, a biosensor assay, a protein chip assay, a mass spectrometry assay, a fluorescence resonance energy transfer assay and an immunostaining assay (*e.g.* immunofluorescence).

20

Standard solid-phase ELISA or FLISA formats are particularly useful in determining the concentration of a protein from a variety of samples.

25

In one form such an assay involves immobilizing a biological sample onto a solid matrix, such as, for example a polystyrene or polycarbonate microwell or dipstick, a membrane, or a glass support (*e.g.* a glass slide). An antibody that specifically binds to a marker of a therapeutic response is brought into direct contact with the immobilized biological sample, and forms a direct bond with any of its target protein present in said

30

sample. This antibody is generally labeled with a detectable reporter molecule, such as for example, a fluorescent label (*e.g.* FITC or Texas Red) or a fluorescent semiconductor nanocrystal (as described in US 6,306,610) in the case of a FLISA or an enzyme (*e.g.* horseradish peroxidase (HRP), alkaline phosphatase (AP) or  $\beta$ -galactosidase) in the case of an ELISA, or alternatively a suitably labeled secondary antibody is used that binds to the first antibody. Following washing to remove any unbound antibody, the label is detected either directly, in the case of a fluorescent label, or through the addition of a substrate, such as for example hydrogen peroxide, TMB, or toluidine, or 5-bromo-4-chloro-3-indol-beta-D-galactopyranoside (x-gal) in the case of an enzymatic label.

Such ELISA or FLISA based systems are suitable for quantification of the amount of a protein in a sample, by calibrating the detection system against known amounts of a protein standard to which the antibody binds.

15

In another form, an ELISA comprises immobilizing an antibody or ligand that specifically binds a marker employed in the prognostic assay of the invention on a solid matrix, such as, for example, a membrane, a polystyrene or polycarbonate microwell, a polystyrene or polycarbonate dipstick or a glass support. A sample is then brought into physical relation with said antibody, and a marker within the polypeptide is bound or 'captured'. The bound protein is then detected using a labeled antibody. For example, if the marker is captured from a human sample, a labeled anti-human antibody that binds to an epitope that is distinct from the first (capture) antibody is used to detect the captured protein. Alternatively, a third labeled antibody can be used that binds the second (detecting) antibody.

25

It will be apparent to the skilled person that the assay formats described herein are amenable to high throughput formats, such as, for example automation of screening processes or a microarray format as described in Mendoza *et al.*, *Biotechniques* 27(4): 778-788, 1999. Furthermore, variations of the above-described assay will be apparent to those skilled in the art, such as, for example, a competitive ELISA.

30

Alternatively, a marker is detected using a radioimmunoassay (RIA). The basic principle of the assay is the use of a radiolabeled antibody or antigen to detect antibody-antigen interactions. An antibody or ligand that specifically binds to the marker is bound to a solid support and a sample brought into direct contact with said antibody. To detect the level of bound antigen, an isolated and/or recombinant form of the antigen is radiolabeled and brought into contact with the same antibody. Following washing, the level of bound radioactivity is detected. As any antigen in the biological sample inhibits binding of the radiolabeled antigen the level of radioactivity detected is inversely proportional to the level of antigen in the sample. Such an assay may be quantitated by using a standard curve using increasing known concentrations of the isolated antigen.

As will be apparent to the skilled artisan, such an assay may be modified to use any reporter molecule, such as, for example, an enzyme or a fluorescent molecule, in place of a radioactive label.

In another example, Western blotting is used to determine the level of a marker employed in the prognostic assay of the invention. In such an assay, protein from a sample is separated using sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) using techniques known in the art and described in, for example, Scopes (*In: Protein Purification: Principles and Practice*, Third Edition, Springer Verlag, 1994). Separated proteins are then transferred to a solid support, such as, for example, a membrane (*e.g.*, a PVDF membrane), using methods known in the art, for example, electrotransfer. This membrane is then blocked and probed with a labeled antibody or ligand that specifically binds to a marker of a therapeutic response. Alternatively, a labeled secondary, or even tertiary, antibody or ligand is used to detect the binding of a specific primary antibody. The level of label is then determined using an assay appropriate for the label used. An appropriate assay will be apparent to the skilled artisan.

For example, the level or presence a protein marker is determined using methods known in the art, such as, for example, densitometry. In one example, the intensity of a protein band or spot is normalized against the total amount of protein loaded on a SDS-PAGE gel using methods known in the art. Alternatively, the level of the marker  
5 detected is normalized against the level of a control/reference protein. Such control proteins are known in the art, and include, for example, actin, glyceraldehyde 3-phosphate dehydrogenase (GAPDH),  $\beta$ 2 microglobulin, hydroxy-methylbilane synthase, hypoxanthine phosphoribosyl-transferase 1 (HPRT), ribosomal protein L13c, succinate dehydrogenase complex subunit A and TATA box binding protein (TBP).

10

In an alternative example, a polypeptide marker of a therapeutic response is detected within a cell, using methods known in the art, such as, for example, immunohistochemistry or immunofluorescence. For example, a cell or tissue section that is to be analyzed to determine the presence of the marker is fixed, to stabilize and  
15 protect both the cell and the proteins contained within the cell. Preferably, the method of fixation does not disrupt or destroy the antigenicity of the marker, thus rendering it undetectable. Methods of fixing a cell are known in the art and include for example, treatment with paraformaldehyde, treatment with alcohol, treatment with acetone, treatment with methanol, treatment with Bouin's fixative and treatment with  
20 glutaraldehyde. Following fixation a cell is incubated with a ligand or antibody capable of binding the marker. The ligand or antibody is, for example, labeled with a detectable marker, such as, for example, a fluorescent label (*e.g.* FITC or Texas Red), a fluorescent semiconductor nanocrystal (as described in US 6,306,610) or an enzyme (*e.g.* horseradish peroxidase (HRP)), alkaline phosphatase (AP) or  $\beta$ -galactosidase.  
25 Alternatively, a second labeled antibody that binds to the first antibody is used to detect the first antibody. Following washing to remove any unbound antibody, the level of the bound to said labeled antibody is detected using the relevant detection means. Means for detecting a fluorescent label will vary depending upon the type of label used and will be apparent to the skilled artisan.

30

Optionally, immunofluorescence or immunohistochemistry will comprise additional steps such as, for example, cell permeabilization (using, for example, n-octyl- $\beta$ D-glucopyranoside, deoxycholate, a non-ionic detergent such as Triton X-100 NP-40, low concentrations of ionic detergents, such as, for example SDS or saponin) and/or antigen  
5 retrieval (using, for example, heat).

Methods using immunofluorescence are preferable, as they are quantitative or at least semi-quantitative. Methods of quantitating the degree of fluorescence of a stained cell are known in the art and described, for example, in Immunohistochemistry (Cuello,  
10 1984 John Wiley and Sons, ASIN 0471900524).

Biosensor devices generally employ an electrode surface in combination with current or impedance measuring elements to be integrated into a device in combination with the assay substrate (such as that described in U.S. Patent No. 5,567,301). An  
15 antibody/ligand that specifically binds to a marker of a therapeutic response is preferably incorporated onto the surface of a biosensor device and a biological sample contacted to said device. A change in the detected current or impedance by the biosensor device indicates protein binding to said antibody. Some forms of biosensors known in the art also rely on surface plasmon resonance to detect protein interactions,  
20 whereby a change in the surface plasmon resonance surface of reflection is indicative of a protein binding to a ligand or antibody (U.S. Patent No. 5,485,277 and 5,492,840).

Biosensors are of particular use in high throughput analysis due to the ease of adapting such systems to micro- or nano-scales. Furthermore, such systems are conveniently  
25 adapted to incorporate several detection reagents, allowing for multiplexing of diagnostic reagents in a single biosensor unit. This permits the simultaneous detection of several proteins or peptides in a small amount of body fluids.

Evanescent biosensors are also preferred as they do not require the pre-treatment of a  
30 biological sample prior to detection of a protein of interest. An evanescent biosensor generally relies upon light of a predetermined wavelength interacting with a fluorescent

molecule, such as for example, a fluorescent antibody attached near the probe's surface, to emit fluorescence at a different wavelength upon binding of the target polypeptide to the antibody or ligand.

- 5 Micro- or nano-cantilever biosensors are also preferred as they do not require the use of a detectable label. A cantilever biosensor utilizes a ligand and/or antibody capable of specifically detecting the analyte of interest that is bound to the surface of a deflectable arm of a micro- or nano-cantilever. Upon binding of the analyte of interest (*e.g.* a marker employed in the prognostic assay of the invention) the deflectable arm of the
- 10 cantilever is deflected in a vertical direction (*i.e.* upwards or downwards). The change in the deflection of the deflectable arm is then detected by any of a variety of methods, such as, for example, atomic force microscopy, a change in oscillation of the deflectable arm or a change in pizoresistivity. Exemplary micro-cantilever sensors are described in USSN 20030010097.

15

- To produce protein chips, the proteins, peptides, polypeptides, antibodies or ligands that are able to bind specific antibodies or proteins of interest are bound to a solid support such as for example glass, polycarbonate, polytetrafluoroethylene, polystyrene, silicon oxide, metal or silicon nitride. This immobilization is either direct (*e.g.* by
- 20 covalent linkage, such as, for example, Schiff's base formation, disulfide linkage, or amide or urea bond formation) or indirect. Methods of generating a protein chip are known in the art and are described in for example U.S. Patent Application No. 20020136821, 20020192654, 20020102617 and U.S. Patent No. 6,391,625. To bind a protein to a solid support it is often necessary to treat the solid support so as to create
- 25 chemically reactive groups on the surface, such as, for example, with an aldehyde-containing silane reagent. Alternatively, an antibody or ligand may be captured on a microfabricated polyacrylamide gel pad and accelerated into the gel using microelectrophoresis as described in, Arenkov *et al. Anal. Biochem.* 278:123-131, 2000.

30

A protein chip may comprise only one protein, ligand or antibody, and be used to screen one or more patient samples for the presence of one polypeptide of interest. Such a chip may also be used to simultaneously screen an array of patient samples for a polypeptide of interest.

5

Preferably, a protein sample to be analyzed using a protein chip is attached to a reporter molecule, such as, for example, a fluorescent molecule, a radioactive molecule, an enzyme, or an antibody that is detectable using methods known in the art. Accordingly, by contacting a protein chip with a labeled sample and subsequent washing to remove  
10 any unbound proteins the presence of a bound protein is detected using methods known in the art, such as, for example, using a DNA microarray reader.

Alternatively, biomolecular interaction analysis-mass spectrometry (BIA-MS) is used to rapidly detect and characterize a protein present in complex biological samples at the  
15 low- to sub-fmole level (Nelson *et al. Electrophoresis* 21: 1155-1163, 2000). One technique useful in the analysis of a protein chip is surface enhanced laser desorption/ionization-time of flight-mass spectrometry (SELDI-TOF-MS) technology to characterize a protein bound to the protein chip. Alternatively, the protein chip is analyzed using ESI as described in U.S. Patent Application 20020139751.

20

As will be apparent from the preceding discussion, it is particularly preferred to employ a detection system that is antibody or ligand based as such assays are amenable to the detection of a marker of a therapeutic response. Immunoassay formats are even more particularly preferred.

25

#### *Biological samples*

As examples of the present invention are based upon detection of a marker in genomic DNA any cell or sample that comprises genomic DNA is useful for determining a disease or disorder and/or a predisposition to a disease or disorder. Preferably, the cell  
30 or sample is derived from a human. Preferably, comprises a nucleated cell.

Preferred biological samples include, for example, whole blood, serum, plasma, peripheral blood mononuclear cells (PBMC), a buffy coat fraction, saliva, a buccal cell, urine, fecal material, sweat, liver biopsy or a skin cell.

- 5 In a preferred example, a biological sample comprises a white blood cell, more preferably, a lymphocyte cell.

Alternatively, the biological sample is a cell isolated using a method selected from the group consisting of amniocentesis, chorionic villus sampling, fetal blood sampling (*e.g.* 10 cordocentesis or percutaneous umbilical blood sampling) and other fetal tissue sampling (*e.g.* fetal skin biopsy). Such biological samples are useful for determining the predisposition of a developing embryo to a therapeutic response.

As will be apparent to the skilled artisan, the size of a biological sample will depend 15 upon the detection means used. For example, an assay, such as, for example, PCR or single nucleotide primer extension may be performed on a sample comprising a single cell, although greater numbers of cells are preferred. Alternative forms of nucleic acid detection may require significantly more cells than a single cell. Furthermore, protein-based assays require sufficient cells to provide sufficient protein for an antigen based 20 assay.

Preferably, the biological sample has been derived or isolated or obtained previously from the subject. Accordingly, the present invention also provides an *ex vivo* method. In one example, the method of the invention additionally comprises isolating, obtaining 25 or providing the biological sample.

In one example, the method is performed using an extract from a biological sample, such as, for example, genomic DNA, mRNA, cDNA or protein.

- 30 As the present invention also includes detection of a marker in a IFN- $\lambda$ 3 gene that is associated with a disease or disorder in a cell (*e.g.* using immunofluorescence), the

term "biological sample" also includes samples that comprise a cell or a plurality of cells, whether processed for analysis or not.

As will be apparent from the preceding description, such an assay may require the use of a suitable control, *e.g.* a normal individual or a typical population, *e.g.*, for quantification.

As used herein, the term "normal individual" shall be taken to mean that the subject is selected on the basis that they are not undergoing treatment with an immunomodulatory composition.

For example, the normal subject has not been diagnosed with any form of medical condition for which therapy may be recommended using, for example, clinical analysis. Alternatively, or in addition, a suitable control sample is a control data set comprising measurements of the marker being assayed for a typical population of subjects known not to suffer from a medical condition for which therapy may be recommended. Preferably the subject is not at risk of developing such a medical condition and *e.g.*, the subject does not have a history of the disease.

In the present context, the term "typical population" with respect to subjects known not to suffer from a disease or disorder and/or comprise or express a marker of a therapeutic response, shall be taken to refer to a population or sample of subjects tested using, for example, known methods for diagnosing the therapeutic response, and determined not to suffer from the disease and/or tested to determine the presence or absence of a marker of the disease, wherein said subjects are representative of the spectrum of normal and/or healthy subjects or subjects known not to suffer from the disease.

In one example, a reference sample is not included in an assay. Instead, a suitable reference sample is derived from an established data set previously generated from a

typical population. Data derived from processing, analyzing and/or assaying a test sample is then compared to data obtained for the sample population.

Data obtained from a sufficiently large number of reference samples so as to be  
5 representative of a population allows the generation of a data set for determining the average level of a particular parameter. Accordingly, the amount of an expression product that is diagnostic of a therapeutic response can be determined for any population of individuals, and for any sample derived from said individual, for subsequent comparison to levels of the expression product determined for a sample  
10 being assayed. Where such normalized data sets are relied upon, internal controls are preferably included in each assay conducted to control for variation.

*Methods for determining a marker associated with therapeutic response*

In another example, the invention additionally comprises determining a marker for a  
15 therapeutic response to any form of medical condition for which therapy with an immunomodulator would be recommended.

Given the tight association between at least one HLA-C allele in combination with at least one IL28B allele and therapeutic response, and the provision of several markers  
20 associated with a therapeutic response, the present invention further provides methods for identifying new markers for a therapeutic response.

Accordingly, the present invention additionally provides a method for identifying a plurality of markers that is associated with a therapeutic response, said method  
25 comprising:

- (i) identifying polymorphisms or alleles or mutations within a HLA-C gene and a IL28B gene or expression products thereof;
- (ii) analyzing a panel of subjects to determine those that suffer from a condition treatable by an immunomodulatory composition and to which the immunomodulatory  
30 composition is administered, wherein not all members of the panel comprise the polymorphisms or alleles or mutations; and

(iii) determining the variation in the development of the therapeutic response to the immunomodulatory composition, wherein said variation indicates that the polymorphisms or alleles or mutations are associated with a subject's response.

- 5 Methods for determining associations are known in the art and reviewed, for example, in King (Ed) Rotter (Ed) and Motulski (Ed), *The Genetic Basis of Common Disease*, Oxford University Press, 2nd Edition, ISBN 0195125827, and Miller and Cronin (Eds), *Genetic Polymorphisms and Susceptibility to Disease*, Taylor and Francis, 1st Edition, ISBN 0748408223.

10

Generally, determining an association between a marker (*e.g.* a polymorphism and/or allele and/or a splice form and/or a mutation) and an event *e.g.*, a response, involves comparing the frequency of a polymorphism, allele, splice form or mutation at a specific locus between a sample of unrelated individuals undergoing treatment (*i.e.*, and  
15 an appropriate control that is representative of the allelic distribution in the normal population.

Several methods are useful for determining associations, however such studies should consider several parameters to avoid difficulties, such as, for example, population  
20 stratification, that may produce false positive results.

Population stratification occurs when there are multiple subgroups with different allele frequencies present within a population. The different underlying allele frequencies in the sampled subgroups may be independent of the disease, disorder and/or phenotype  
25 within each group, and, as a consequence, may produce erroneous conclusions of linkage disequilibrium or association.

Generally, problems of population stratification are avoided by using appropriate control samples. For example, case-comparison based design may be used in which a  
30 comparison between a group of unrelated probands with the disease, disorder and/or phenotype and a group of control (comparison) individuals who are unrelated to each

other or to the probands, but who have been matched to the proband group on relevant variable (other than infection status) that may influence genotype (e.g. sex, ethnicity and/or age).

- 5 Alternatively, controls are screened to exclude those subjects that have a personal history of a disease or treatment. Such a "supernormal" control group is representative of the allele distribution of individuals unaffected by a disease or treatment.

In general, an analysis of association is used to detect non-random distribution of one  
10 or more alleles and/or polymorphisms and/or splice variants within subjects affected by a disease/disorder and/or phenotype of interest. The comparison between the test population and a suitable control population is made under the null hypothesis assumption that the locus to which the alleles and/or polymorphisms are linked has no influence on phenotype, and from this a nominal p-value is produced. For analysis of a  
15 biallelic polymorphism or mutation (e.g. a SNP) using a case control study, a chi-square analysis (or equivalent test) of a 2 x 2 contingency table (for analysis of alleles) or a 3 x 2 contingency table (for analysis of genotypes) is used.

For analysis using a family-based association study, marker data from members of the  
20 family of each proband are used to estimate the expected null distributions and an appropriate statistical test performed that compares observed data with that expected under the null hypothesis.

Another method useful in the analysis of association of a marker with a disease,  
25 disorder and/or phenotype is the genomic control method (Devlin and Roeder, *Biometrics*, 55: 997-1004, 1999). For a case-control analysis of candidate allele/polymorphism the genetic control method computes chi-square test statistics for both null and candidate loci. The variability and/or magnitude of the test statistics observed for the null loci are increased if population stratification and/or unmeasured  
30 genetic relationships among the subjects exist. This data is then used to derive a multiplier that is used to adjust the critical value for significance test for candidate loci.

In this manner, genetic control permits analysis of stratified case-control data without an increased rate of false positives.

A structured association approach (Pritchard *et al.*, *Am. J. Hum. Genet.*, 67: 170-181, 2000) uses marker loci unlinked to a candidate marker to infer subpopulation membership. Latent class analysis is used to control for the effect of population substructure. Essentially, null loci are used to estimate the number of subpopulations and the probability of a subject's membership to each subpopulation. This method is then capable of accounting for a change in allele/polymorphism frequency as a result of population substructure.

Alternatively, or in addition, a Bayesian statistical approach may be used to determine the significance of an association between an allele and/or polymorphism in a gene and a response to treatment. Such an approach takes account of the prior probability that the locus under examination is involved in the therapeutic response of interest (e.g., Morris *et al.*, *Am. J. Hum. Genet.*, 67: 155-169, 2001).

Publicly available software may be employed to determine associations.

The present invention is further described with referenced to the following non-limiting examples.

#### Example 1

Identification of alleles in IL28B associated with therapy of chronic hepatitis C using an immunomodulatory composition comprising an interferon (IFN)

#### Summary

This example demonstrates SNPs and alleles of the present invention that are associated with a response to therapy for hepatitis C virus infection using a composition comprising IFN. The disclosed associations between variations in the IL28B gene are

sufficiently-strong to indicate that genotypes in 19q13.13 between position 44,423,000 and position 44,436,000, especially IFN- $\lambda$ 3 (IL-28B) genotypes, can be used to predict drug responses. The data also support the use of IFN- $\lambda$  *e.g.*, IFN- $\lambda$ 1 and/or IFN- $\lambda$ 2 and/or IFN- $\lambda$ 3, for treatment of HCV infection and other diseases currently treated  
5 using other IFNs such as IFN- $\alpha$  or IFN- $\beta$  or combinations thereof.

For example, the data herein provide high response (HR) alleles associated with a rapid or strong or sustained virus response (SVR) of subjects to hepatitis C therapy comprising IFN- $\alpha$  and ribavirin, as determined by virus clearance, in addition to  
10 providing low response (LR) alleles associated with a poor or low response or no response to such therapy. Tables 1 and 3 provide listings of SNPs linked to the IL28B gene on chromosome 19 which form part of the applicants' invention, including specific examples of HR alleles and LR alleles that are predictive of the response to treatment.

15 Table 4 provides data on the effects of different genotypes comprising exemplary HR alleles and/or LR alleles linked to IL28B including, for example, HR alleles and/or LR alleles having approximate p value less than about  $10^{-9}$  for rs8099917, rs8109886, rs10853727, rs8103142 and rs12980275, in addition to several weak alleles of rs4803224, rs12980602 and rs10853728. Table 5 provides data on the effects of  
20 different genotype combinations comprising exemplary HR alleles and/or LR alleles linked to IL28B, including, for example, rs8099917, rs8103142, and rs12980275. For example, double-homozygotes for the HR alleles of rs12980275 and rs8099917, and double-homozygotes for the HR alleles of rs8103142 and rs8099917 and triple homozygotes for the HR alleles of rs12980275, rs8103142 and rs8099917 show strong  
25 responses to therapy ( $p < 6 \times 10^{-4}$ ), whereas the corresponding homozygotes for LR alleles at these loci demonstrate consistently low responses to therapy ( $p > 0.04$ ), as shown in Table 5.

Table 6 provides data on haplotype effects for haplotypes comprising exemplary HR  
30 alleles and/or LR alleles linked to IL28B including, for example, the haplotype

comprising rs12980275, rs8105790, rs8103142, rs10853727, rs8109886 and rs8099917.

Table 7 provides data on the sensitivity, specificity, and positive predictive value (PPV) of different genotypes comprising exemplary HR alleles and/or LR alleles linked to IL28B. These data demonstrate that two doses of a LR allele *i.e.*, homozygosity of the LR allele, is highly predictive of a poor, weak or non-response to therapy, especially with respect to LR alleles of rs7248668, rs4803221, rs8099917 and rs12979860. By virtue of the significant linkage disequilibrium between markers linked to IL28B, the same conclusions may be drawn in respect of HR alleles and/or LR alleles which are inherited as haplotype blocks with one or more of rs7248668, rs4803221, rs8099917 and rs12979860. The data in Table 7 also demonstrate that homozygosity of a LR allele of rs7248668 and/or homozygosity of a LR allele of rs4803221 provide(s) a higher PPV than rs8099917, which is the currently-accepted "gold standard" for determining a likelihood that a subject will respond well to treatment, suggesting a lower incidence of false positives in testing procedures based on homozygosity of LR alleles of rs7248668 and/or homozygosity of LR alleles of rs4803221 than for tests requiring rs8099917 genotyping. Thus, the data herein indicate that the HR allele and/or the LR allele of rs7248668 and/or the HR allele and/or the LR allele of rs4803221 provide improved prognostic value than LR and/or HR alleles of rs8099917.

These data provide the means for predicting outcome of therapy to immunomodulatory compositions with accuracy in more than 80% or 81% or 82% or 83% or 84% or 85% or 86% or 87% or 88% or 89% or 90% of cases.

25

#### *Genome-Wide Association Study (GWAS)*

##### Patient cohorts

For stage one genotyping, a well-characterised Australian population of 302 patients of European ancestry, matched for age, BMI, viral titre, and treatment regime was employed (Table 2). Patients were excluded from the study if they had been co-infected with either HBV or HIV or if they were not of European descent. All patients

included in this study had been diagnosed as infected with genotype 1 HCV based on serology and viral DNA tests, had received a standard course of pegylated interferon-alpha (IFN- $\alpha$ ) and ribavirin, and their six-month post-treatment responses to therapy as determined by virus clearance had been determined. All patients who responded to therapy, and most patients classified as having a non-sustained viral response ("non-SVR"), had received treatment for 48 weeks. A few non-SVR cases received only 24 weeks of therapy, because they showed no reduction in viral RNA at week 12. All patients were seen by experienced hepatologists at their respective hospitals.

10 A larger independent cohort consisting of about 600 Europeans from the United Kingdom, Germany, Italy and Australia was also employed for stage two genotyping (Table 2). The criteria for recruitment of study subjects in this cohort were the same for the Australian cohort.

15 Sample collection and processing

Australian samples for both stages were collected at Sydney (Westmead Hospital, Nepean Hospital, St Vincent's Hospital and Prince of Wales Hospital) and Brisbane (Princess Alexandra Hospital). Case samples for the replication cohorts were collected at Universtat Zu Berlin, Germany (n=298), Rheinische Friedrich-Wilhelms-Universitat, Bonn, Germany (n=43), Universita degli Studi di Torino, Turin, Italy (n=93) and Freeman Hospital, Newcastle, UK (n=91).

Blood was collected into EDTA tubes (Australian cohort). Extracted DNA normalised to 50 ng/ul was obtained for other cohorts. Genomic DNA was extracted by standard protocols. DNA quality was assessed by calculating absorbance ratio OD<sub>260 nm</sub>/280 nm using nanodrop.

Ethics approval

Ethical approval for this study was given by Sydney West Area Health Service Human Research Ethics Committee at Westmead Hospital and the University of Sydney (HREC No.2002/12/4.9 (1564)). All other sites had ethical approval from their

respective ethics committees. Written informed consent was obtained from all participants.

#### Statistical analysis

5 Hardy-Weinberg equilibrium and allelic distributions in subjects having high response(s) or low response(s) were compared using a chi-squared test in Haploview version 3.31 of the Broad Institute, USA *e.g.*, as described by Barrett *et al. Bioinformatics* 21, 263-265 (2005). The threshold for genome-wide significant association was set at  $p < 1.6 \times 10^{-7}$  *i.e.*,  $0.05/312,000$ . SNPs having  $1.6 \times 10^{-7} \leq p < 1.0 \times 10^{-4}$  were considered to show a highly suggestive association with response to therapy. SNPs having  $1.0 \times 10^{-4} \leq p < 1.0 \times 10^{-3}$  were considered to show a moderately suggestive association with response to therapy. The Cochran-Armitage trend test was used to assess association of all SNPs tested in Stage one and Stage two, and merged p values were determined.

15

#### SNP genotyping

A two-stage approach essentially as described by Saito *et al. J. Hum. Genetics* 47, 360-365 (2002) was employed for SNP genotyping using HumanLinkage panels for Infinium and GoldenGate SNP Genotyping (Illumina, Inc., San Diego, USA). SNPs on 20 Chr 19 were fine-mapped using the Sequenom mass array iPlex genotyping platform (Sequenome, Inc, San Diego, USA). The two-stage approach was favoured as it was calculated to have a power of 87% to detect risk factors of 1.5 for disease allele frequency of 0.2 (Skol *et al, Nature Genetics* 38, 209-213 (2006).

#### 25 Stage one genotyping results

The 302 patient samples were genotyped using the Infinium HumanHap300 or CNV370 genotyping BeadChip (Illumina, Inc., San Diego, USA). Samples having a very low call rate using the Illumina cluster (*i.e.*, genotyping efficiency less than 95%) was deleted. A minor allele frequency (MAF) check was performed for data handling 30 accuracy, and those SNPs occurring in less than 0.05% of samples were deleted. Samples providing a Hardy-Weinberg equilibrium p value  $> 10^{-4}$  were retained. Two

individuals were excluded due to genotyping call rates less than 90%, and IBS/IBD analysis revealed that those two individuals were related. The 99% confidence interval (CI) for genotyping error was estimated to be between 1.7% and 1.8%. As ethnicity was determined by self-identification or parental ethnic identification, an assessment  
5 for possible population stratification was performed using EIGENSOFT software, essentially as described by Price *et al. Nature Genetics* 38, 904-909 (2006), applying principal component analysis to the genotype data to infer the axes of variation. This resulted in exclusion of five individuals from further analysis. Accordingly, the final genome-wide association (GWA) study consisted of 293 patients (162 having low  
10 response(s) (LR) and 131 having high response(s) (HR) as indicated in Table 2.

A Manhattan plot of signal intensity relative to genome position and a Quantile-Quantile plot of allelic associations for the stage one SNPs (not shown), identified SNPs that were more associated than may be expected by mere chance. A genomic  
15 inflation factor lambda of 1.005 in that analysis indicated a low possibility of false positive associations *e.g.*, due to population stratification. A total of 312,000 SNPs passed the first stage quality filters and were analysed further. A total of 695 SNPs (0.22%) did not pass the first stage quality filters and were excluded from subsequent analysis.

20

From these 312,000 SNPs, SNPs were classified as highly or suggestively associated with the therapeutic response, as described under "statistical analysis" *supra*.

In stage one, three chromosome 19 SNPs were identified having a high or suggestive  
25 association with therapeutic response that were linked to the interferon lambda-3 (IFN- $\lambda$ 3) gene. No other SNPs mapping to chromosome 19 satisfied the threshold for genome-wide associations. Two SNPs flanking the IFN- $\lambda$ 3 gene *i.e.*, rs8099917 mapping to the 5'-end of IFN- $\lambda$ 3 ( $p=7.06 \times 10^{-08}$ ), and rs12980275 mapping to the 3'-end of IFN- $\lambda$ 3 ( $p=4.81 \times 10^{-8}$ ) were well-below the threshold for significant  
30 associations with therapeutic response in stage one. The third SNP *i.e.*, rs8109886

mapping to the 5'-end of IFN- $\lambda$ 3 ( $p=1.29 \times 10^{-04}$ ) was considered to have a suggestive association with therapeutic response in stage one.

SNPs on other chromosomes (not shown) were also identified having at least  
5 moderately suggestive positive associations with therapeutic response that mapped to the following chromosomal locations:

- a) rs7512595 at about 1p35, between WASF2 and ADHC1 genes;
- b) rs6806020 between about 3p21.2 and about 3p21.31, within an intron of the CACNA2D3 gene; and rs12486361 between about 3p24.3 and about 3p25.1,  
10 within an intron of the RTFN-1 gene;
- c) rs10018218 at about 4q32; rs1581096 at about 4p13; and rs1250105 at about 4p16.1 near to the CTBP1 gene;
- d) rs7750468 at about 6q22.31, between C6orf68 and SLC35F1; rs2746200 at about 6q13, within an intron of the RIMS-1 gene; rs927188 between about  
15 6p12.2 and about 6p12.3, within an intron of PHKD-1; rs2517861 between about 6p21.33 and about 6p22 and between HLA pseudogenes HCP5P10 and MICF; and rs2025503 and rs2066911 between about 6p22.1 and about 6p22.2, and between ALDH5A1 and PRL genes;
- e) rs10283103 and rs2114487 between about 8q12.2 to about 8q13.1 *e.g.*, between  
20 CRH and MGC33510 such as between ADHFE1 and MGC33510 or between RRS1 and CRH;
- f) rs1002960 between about 9q22.1 and about 9q22.2, in an intergenic region;
- g) rs1931704 between about 10q26.2 and about 10q26.3 and between NPS and DOCK1;
- 25 h) rs1939565 at about 11q21, between KIAA1731 and FN5 genes; rs568910 and rs557905 within introns of the CASP-1 gene, wherein rs568910 is in intron 2 and rs557905 is in intron 6;
- i) rs1931704 between about 14q22.1 and 14q22.2, between DACT1 and LOC729646;

- j) rs3093390 between about 16p11.2 and about 16p12.1 and between IL21R and GFT3C1; and rs7196702 between about 16q23.1 and about 16q23.2, in an intron of the WWOX gene; and
- k) rs4402825, between about 20q13.12 and about 20q13.13 in an intron of the SULF2 gene.

The SNPs, rs7750468, rs2066911, rsrs6806020, rs2114487 and rs1931704 were considered highly suggestive of an association, and the remaining considered to be moderately suggestive of an association based on stage one screening data. The SNP designated rs1931704 has a very close association with therapeutic response ( $p=4.42 \times 10^{-07}$ ), and was shown to be closely-linked to the neuropeptide S (NPS) gene.

A total of 512 highly and moderately associated SNPs were selected from stage one.

#### 15 Stage two genotyping results

In the second stage whole genome screen, 307 SNPs having a significance level of  $p \leq 1.0 \times 10^{-4}$  irrespective of their genome location, and 206 SNPs linked to genes classified as immune regulatory or anti-viral by gene ontology and having a significance level of  $1.0 \times 10^{-4} \leq p < 1.0 \times 10^{-3}$  were included. The SNPs were genotyped using Golden Gate technology (Illumina, Inc., San Diego, USA). Two (2) cases having call rates of less than 0.90, 8 samples with no treatment outcome were excluded. Cluster plots of the remaining samples were checked by visual inspection and 38 ambiguous calls and SNPs with MAF less than 0.05 were also excluded from further analysis. A further 8 SNPs were excluded as having poor significance in their Hardy-Weinberg equilibrium *i.e.*,  $p$  value  $< 10^{-4}$ . This meant that the stage two analysis was carried out in 577 individuals, of which 294 had low response(s) and 261 had high response(s) to therapy (Table 2).

A total of 468 SNPs passed the quality filters and were selected for stage 2 genotyping. These 468 SNPs were classified as highly or suggestively associated with the therapeutic response, as described under "statistical analysis" *supra*. Of these, in stage

2, 40 SNPs achieved the threshold for suggestive evidence of association with treatment response in the replication phase ( $p \leq 0.05$ ).

As shown in Table 4, SNPs linked to the IFN- $\lambda$ 1 (IL28B) gene showed moderate-to-  
5 strong associations with therapeutic response, including rs8099917 in the 5'-end of the gene, which provided the most highly-significant association ( $p = 9.39 \times 10^{-04}$ ; OR=1.56; and 95% CI=1.19-2.04). A moderate association was observed for rs12980275 mapping to the 3'-end of IFN- $\lambda$ 3 ( $p = 1.24 \times 10^{-4}$ ), which had provided a higher significance value in the previous cohort (Table 4). Also shown in Table 4, moderate  
10 associations were also observed for the SNPs rs8103142 in exon 2 of the IL28B gene ( $p = 3.83 \times 10^{-4}$ ) and rs8105790 in the 3'-end of the IL28B gene ( $p = 3.7 \times 10^{-4}$ ).

Associations with therapeutic outcome were weaker in the stage two cohort for SNPs mapped to other genomic regions, with the exception of rs10018218 and rs1002960,  
15 which provided moderate associations.

#### Merged data

The Cochran-Armitage trend test (Cochran *Biometrics* 10, 417-451, 1954; Armitage *Biometrics* 11, 375-386, 1955) was used to assess association of all SNPs tested in  
20 stage one and stage two, and merged P values were determined.

Data presented in Table 4 reveal strong associations with response, reaching genome-wide significance in the overall analysis of the discovery and replication groups, for rs8099917, and rs12980275 flanking the IL28B gene on chromosome 19, with a strong  
25 association for rs8109886 in the 5'-end of the IL28B gene.

#### SNP genotyping to determine high response (HR) and low response (LR) alleles

Conventional methods are used to determine genotypes for the various SNPs listed in Tables 1, 3 and 4, such as, for example a method selected from the following and  
30 combinations and variations thereof:

- (i) by hybridizing complementary DNA probes to the SNP site in genomic DNA *e.g.*, under high stringency hybridization conditions;
- (ii) by dynamic allele-specific hybridization (DASH) employing fluorescently-labelled allele specific oligonucleotides to hybridize to single-stranded biotinylated  
5 genomic DNA amplicons bound to a streptavidin column;
- (iii) by using a molecular beacon such comprising a sequence of a wild-type allele or a mutant allele;
- (iv) by interrogating a SNP microarray using probes comprising the SNP site in several different locations or comprising mismatches to the SNP allele and comparing  
10 the signal intensities produced to thereby determine homozygous and heterozygous alleles;
- (v) by analyzing restriction fragment length polymorphisms (RFLPs) generated by digestion of genomic DNA using enzymes that distinguish sequence comprising a SNP and resolution of the fragments produced based on their lengths;
- 15 (vi) by PCR or other amplification means employing *e.g.*, ARMS primers of different length or differentially labelled and comprising sequences that overlap at the SNP site to thereby amplify the alleles;
- (vii) by Invader assay employing *e.g.*, a flap endonuclease (FEN) such as cleavase to digest a tripartite structure comprising genomic DNA and two specific oligonucleotide  
20 probes wherein a first probe (the Invader oligonucleotide) is complementary to the 3' end of the genomic DNA and comprises a mismatched 3'-terminal nucleotide that overlaps the SNP in the target genomic DNA and wherein a second probe (allele-specific oligonucleotide) is complementary to the 5' end of the target genomic DNA and extends past the 3' side of the SNP nucleotide and comprises a nucleotide  
25 complementary to a SNP allele, such that the tripartite structure forms when the SNP is in the target genomic DNA and cleavase releases the 3' end of the allele-specific probe from the tripartite structure when the matched allele is present in the allele-specific oligonucleotide;
- (vii) by primer extension across the SNP from a probe that is hybridized to the  
30 genomic DNA immediately upstream of the SNP nucleotide in the presence of mixes of

- dNTP/ddNTP mixes each lacking a different ddNTP and sequencing the extension products produced;
- (viii) by iPLEX SNP genotyping (Sequenom Inc., San Diego, USA);
  - (ix) by arrayed primer extension (APEX or APEX-2);
  - 5 (x) by Infinium assay (Illumina Inc., San Diego, USA) based on primer extension;
  - (xi) by homogeneous multiplex PCR employing two oligonucleotide primers per SNP to generate amplicons that comprise the alleles in genomic DNA;
  - (xii) by 5'-nuclease assay employing a thermostable DNA polymerase having 5'-nuclease activity to degrade genomic DNA hybridizing to matched primers but not
  - 10 mismatched primers *e.g.*, performed in real time such as in a Taqman assay format (Applied Biosystems, Carlsbad, USA) and/or in a multiplex assay format;
  - (xiii) by ligase assay employing matched and mismatched oligonucleotides to interrogate a SNP by hybridizing the probes over the SNP site such that ligation to an upstream or downstream constant oligonucleotide can occur if the probes are identical
  - 15 to the target genomic DNA;
  - (xiv) by analyzing single strand conformation polymorphisms *e.g.*, as determined by mobility of single-stranded genomic DNA or amplicons produced therefrom;
  - (xv) by temperature gradient gel electrophoresis (TGGE) or temperature gradient capillary electrophoresis (TGCE) employing target DNA comprising denaturing target
  - 20 DNA comprising the SNP site in the presence of an allele-specific probe comprising a mismatched allele to the target DNA, re-annealing the nucleic acids and resolving the products in the presence of a temperature gradient;
  - (xvi) by denaturing HPLC comprising denaturing target DNA comprising the SNP site in the presence of an allele-specific probe comprising a mismatched allele to the
  - 25 target DNA, re-annealing the nucleic acids and resolving the products under reverse-phase HPLC conditions;
  - (xvii) by high-resolution melting of amplicons;
  - (xviii) by SNPLex (Applied Biosystems, Carlsbad, USA); and
  - (xix) by sequencing across SNPs in genomic DNA *e.g.*, employing pyrosequencing.

For example, using conventional methods, HR and LR alleles were determined for SNPs exemplified herein, and these are summarized in Table 3 hereof in the column headed "SNP effect".

5 In one particular example, the rs8099917 SNP was typed by PCR-RFLP using *Tsp45I* restriction enzyme (New England Biolabs, Beverley, MA). Digestions were performed in 10 µl reactions in X1 buffer, 0.4U enzyme, 5 µl PCR product and Milli Q water at 65°C for 2h. Digested products were electrophoresed at 120V for 1/2 h on a 2% (w/v) TBE gel. Genotype was determined as a 325 bp fragment for the T allele, and as  
10 fragments of 286 bp and 39 bp for the G allele. Release of a further 214bp fragment arising from digestion at an internal control *Tsp45I* site was used to assess completeness of digestion. Data in Table 4 show a very strong association with therapeutic response reaching genome-wide significance for the T allele (and complementary A residue on the opposing DNA strand) of rs8099917 (merged P value  
15 =  $9.25 \times 10^{-09}$ , OR=1.86, 95% CI= 1.49-2.32). Compared to non-carriers of the high response (HR) allele at rs8099917, heterozygous carriers of the rs8099917 HR allele produced an odds ratio (OR) of 1.64 (95% CI = 1.15-2.32) and homozygous carriers produced an OR of 2.39 (95% CI = 1.16 – 4.94).

20 Data in Table 4 also show a very strong association with therapeutic response reaching genome-wide significance for the A allele (and complementary T residue on the opposing DNA strand) of rs12980275 (merged P value =  $7.74 \times 10^{-10}$ ).

Data in Table 4 also indicate that the T allele (and complementary A residue on the  
25 opposing DNA strand) of rs8103142 in exon 2 of IL28B is associated with a higher response to therapeutic intervention with immunomodulatory compositions *i.e.*, it is the HR allele ( $p=3.83 \times 10^{-4}$ ), whereas the C allele (and complementary G residue on the opposing DNA strand) are associated with a lower response *i.e.*, it is the low response (LR) allele.

Data presented in Table 5 show the possible genotypes for two-SNP and three-SNP combinations comprising rs12980275, rs8099917 and rs8103142. These data support the use of combinations of HR alleles and/or LR alleles for the individual SNPs. For example, there is a highly-significant association between therapeutic response and the combination of both HR alleles in rs12980275 and rs8099917 *i.e.*, genotype AA at rs12980275 and genotype TT at rs8099917 ( $p = 6.13 \times 10^{-5}$ ; OR= 2.11; 95% CI= 1.46-3.04). Similarly, there is a highly-significant association between therapeutic response and the combination of both HR alleles in rs8103142 and rs8099917 *i.e.*, genotype TT at rs8103142 and genotype TT at rs8099917 ( $p = 4.92 \times 10^{-4}$ ; OR= 2.03; 95% CI= 1.36-3.05). The triple homozygotes for HR alleles at these loci are also associated at high significance with therapeutic response *i.e.*, genotype AA at rs12980275 and genotype TT at rs8103142 and genotype TT at rs8099917 ( $p = 6.3 \times 10^{-4}$ ; OR= 2.03; 95% CI= 1.36-3.05).

Collectively, the data presented in Tables 4 and 5 suggest that genotypes in 19q13.13 between position 44,420,000 and position 44,440,000 and more specifically between about position 44,423,000 and about position 44,436,000, especially IFN- $\lambda$ 3 (IL-28B) genotypes, are predictive of patient responses to immunomodulatory compositions *e.g.*, an interferon such as IFN- $\alpha$  and/or an agent that modulates Th1/Th2 such as ribavirin.

#### Haplotype analysis for IFN- $\lambda$ 3 (IL28B)

Haplotypes of SNPs linked to the IFN- $\lambda$ 3 (IL28B) gene were selected using Haplotype Tagger software of the Center for Human Genetic Research of Massachusetts General Hospital and Harvard Medical School, USA, and the Broad Institute, USA, *e.g.*, as described by de Bakker *et al.*, *Nature Genetics* 37, 1217-1223 (2005). Haplotype Tagger is a tool for the selection and evaluation of SNPs from genotype data, that combines a pairwise tagging method with a multimarker haplotype approach. Genotype data and/or a chromosomal location within which SNPs are mapped are provided as a source for calculation of linkage disequilibrium patterns based on sequence data for the chromosomal region of interest. Haplotype Tagger provides a list of the SNPs and corresponding statistical tests that capture variants of interest.

Haplotype Tagger may be implemented in the stand-alone program, Haploview (*e.g.*, version 3.31) of the Broad Institute, USA (*e.g.*, Barrett *et al. Bioinformatics* 21, 263-265, 2005).

- 5 In one example, Haplotype Tagger was employed to fine-map SNPs in the IFN- $\lambda$  gene cluster and to tag the common haplotypes in the chromosomal region comprising the IFN $\lambda$  gene cluster. That analysis identified IFN- $\lambda$ 3 as having a distinct haplotype block for alleles at loci identified herein as being associated with therapeutic response (data not shown).

10

Pairwise correlation coefficients were determined for IFN- $\lambda$ 3 SNPs and haplotype distributions within the study population were determined. For example, Table 6 shows haplotypes for combinations of the following SNPs:

- (a) rs12980275, for which possible alleles are A or G (SEQ ID NO: 87);  
 15 (b) rs8105790 for which possible alleles are C or T (SEQ ID NO: 84);  
 (c) rs8103142 for which possible alleles are C or T (SEQ ID NO: 66);  
 (d) rs10853727 for which possible alleles are C or T (SEQ ID NO: 31);  
 (e) rs8109886 for which possible alleles are A or C (SEQ ID NO: 9); and  
 (f) rs8099917 for which possible alleles are G or T (SEQ ID NO: 4).

20

The data presented in Table 6 show that the G allele for rs8099917 *i.e.*, the LR allele, tags the haplotype that is most-associated with a low response to therapy ( $p=3.03 \times 10^{-9}$ ; OR=2.0; 95% CI=1.58-2.50).

- 25 The data presented in Table 6 also show that the HR allele for rs12980275 is linked to HR alleles at rs8105790, rs8103142, rs8109886 and rs8099917 in 45.2% of the test population, and that the LR allele for rs8099917 is associated with LR alleles for rs12980275, rs8105790, rs8103142, rs10853727 and rs8109886 in 25.6% of the test population, suggesting linkage disequilibrium between these alleles. Accordingly, the  
 30 occurrence of specific alleles linked to IFN- $\lambda$ 3 may predict haplotypes associated with high or low responses to therapy.

*Massively-parallel sequencing and Burrows-Wheeler Alignments*Methods

In another approach to identify SNPs or indels in the IL28 region of human chromosome 19 having predictive capacity in determining a likelihood that chronically-  
5 infected HCV-infected subjects will respond to treatment with IFN- $\alpha$  plus ribavirin, massively parallel sequencing of the IL28 genomic region was performed on pools of genomic DNA from 100 subjects each, that were derived from either (i) HCV-infected subjects that responded to treatment with IFN- $\alpha$  plus ribavirin (HR) or (ii) subjects having a poor response or no response to treatment with IFN- $\alpha$  plus ribavirin (LR).  
10 Briefly, paired end reads were collected on an Illumina GAII sequencer. The data were analyzed *in silico* to thereby identify genetic variants that were represented differently in the genomes of the two pools of subjects. For example, the Burrows-Wheeler Alignment (BWA1) was employed to perform short-read alignments of the sequences of the genetic variants *e.g.*, as described by Li *et al.*, *Bioinformatics*, 25, 1754-1760  
15 (2009). BWA1 permits gapped alignments to thereby facilitate a detection of indels, however does not use quality scores to determine alignment locations, but maps reads from paired end data as singletons if no residue alignment is otherwise determined. This identifies all genetic variations between the two cohorts that potentially explain the variation in response to treatment.

20

Non-parametric measures of the statistical dependencies between known SNPs linked to the IL28B region of the genome according to HapMap and the genetic variations obtained by massively parallel sequencing were determined as Spearman Correlation Coefficients. This confirms that the genetic variability is correlated to HR and LR  
25 alleles of SNPs linked to IL28B.

The genotypes of about 460 individuals infected with HCV and having known treatment responses were also determined, and the sensitivities, specificities, and positive predictive values (PPVs), and negative predictive values (NPVs) were  
30 calculated based on the frequencies of those genotypes, for several SNPs linked to IL28B, as follows:

a) For LR alleles:

1. Sensitivity was determined as the number of individuals that are homozygous for a test LR allele of a SNP linked to IL28B and do not respond significantly to treatment and/or do not have a sustained viral response to treatment (*i.e.*, low-responders and non-responders), as a proportion of the number of all low-responders and non-responders in the same cohort;
2. Specificity was determined as the number of individuals that are *not* homozygous for a test LR allele of a SNP linked to IL28B and respond significantly to treatment and/or have a sustained viral response (*i.e.*, responders and high-responders), as a proportion of the total number of responders and high-responders in the same cohort;
3. PPV was determined as the number of low-responders and non-responders that are homozygous for a test LR allele of a SNP linked to IL28B as a proportion of the number of all subjects in the same cohort that are also homozygous for the same test LR allele; and
4. NPV was determined as the number of responders and high-responders that are *not* homozygous for a test LR allele of a SNP linked to IL28B as a proportion of the number of all subjects in the same cohort that are also *not* homozygous for the same test LR allele.

b) For HR alleles:

1. Sensitivity was determined as the number of responders and high-responders that are homozygous for a test HR allele of a SNP linked to IL28B as a proportion of the number of all responders and high-responders in the same cohort;
2. Specificity was determined as the number of individuals that are *not* homozygous for a test HR allele of a SNP linked to IL28B as a proportion of the number of all low-responders and non-responders in the same cohort;
3. PPV was determined as the number of responders and high-responders that are homozygous for a test HR allele of a SNP linked to IL28B as a proportion of the

number of all subjects in the same cohort that are also homozygous for the same test HR allele; and

4. NPV was determined as the number of low-responders and non-responders that are *not* homozygous for a test HR allele of a SNP linked to IL28B and do not  
5 respond significantly to treatment as a proportion of the number of subjects in the same cohort that are also *not* homozygous for the same test HR allele.

Thus, the positive predictive value was determined to measure the proportion of subjects who are correctly diagnosed by determining homozygosity at HR and LR  
10 alleles linked to IL28B.

### Results

Massively parallel sequencing has confirmed the identity of SNPs identified by GWAS herein, and also identified rs7248668 as being associated with the response of subjects  
15 to therapy, including the increased likelihood that subjects who are homozygous for LR alleles of rs7248668 or rs4803221 will not mount a significant and/or sustained viral response (SVR) to therapy, and/or the increased likelihood that subjects who are homozygous for HR alleles of rs7248668 or rs4803221 will mount a significant and/or sustained viral response (SVR) to therapy. The SNPs rs7248668 and rs4803221 are in  
20 linkage disequilibrium with SNPs identified by GWAS, including rs8099917 and rs12979860, *e.g.*, a correlation coefficient ( $r^2$ ) of 0.96 for responders and high-responders (HR) and a correlation coefficient ( $r^2$ ) of 0.97 for poor-responders and non-responders (LR).

- 25 The data provided in Table 7 indicate that LR alleles of the SNPs designated rs4803221 and rs7248668 and rs8099917 and rs12979860 provide accurate predictions of low-response or non-response to anti-HCV therapy, as shown by a PPV of greater than 70% in each case for two copies of the LR allele. This provides the first statistical ranking of LR alleles, and demonstrates for the first time the improved prognostic value of LR  
30 alleles at rs4803221 and rs7248668 relative to rs8099917.

If the minor (LR) allele of rs4803221 or rs7248668 is used to determine treatment response without considering contributions from other markers, the test may be expected to provide 88.5% accuracy for homozygotes as each locus, as determined by the PPV (Table 7). In contrast, the minor (LR) allele of rs8099917 may provide only  
5 81.8% accuracy for homozygotes as this locus, and the minor (LR) allele of rs12979860 may provide only 74.4% accuracy for homozygotes as this locus, as determined by the PPV (Table 7).

In all cases, the population coverage of each of these minor alleles as a stand-alone test  
10 is quite low. However, tests for homozygosity of LR alleles at rs4803221 or rs7248668 have slightly improved specificity relative to tests for homozygosity of LR alleles at rs8099917 or rs12979860, indicating that the markers can be combined without substantially compromising test specificity, and/or test PPV.

15 Testing for combinations of markers, such as combinations of homozygous LR alleles at any two or more of rs4803221 and/or rs7248668 and/or rs8099917 and/or rs12979860, will also improve test sensitivity in the population, by providing wider population coverage than single markers. For example, by testing for the combination of rs4803221 and rs8099917, population coverage is nearly doubled, and by testing for  
20 the combination of rs4803221 and rs8099917 and rs12979860, up to approximately 33.4% of the population may be covered. Similarly, by testing for homozygosity of LR alleles at rs7248668, only about 7.8% of the population is covered, whereas testing for rs7248668 in combination with homozygosity of LR alleles at rs4803221 provides nearly double that population coverage, and testing for homozygosity of LR alleles at  
25 rs7248668 and rs4803221 and rs8099917 provides about 21.8% population coverage, whereas testing for homozygosity of LR alleles at rs7248668 and rs4803221 and rs8099917 and rs12979860 provides about 41% population coverage.

Two SNPs provide better prediction than rs8099917 and rs1297860 (Table 1):  
30 rs4803221 and rs7248668. The new SNPs are in linkage disequilibrium with the latter SNPs, but tag response slightly better than the previous best SNPs.

*Determining expression of IFN $\lambda$ -3 (IL28B)*

Total RNA was extracted from whole blood cells of healthy controls according to standard procedures, and used as a template for single-stranded cDNA synthesis using random hexamer primers and Superscript III reverse transcriptase (Invitrogen) according to manufacturer's instruction. RT-PCR was performed employing primers and probes for IFN- $\lambda$ 1 (IL29), IFN- $\lambda$ 2 (IL28A) and IFN- $\lambda$ 3 (IL28B), essentially as described by Mihm *et al.*, *C. Lab. Invest.* 84, 1148-1159 (2004). The expression levels of the mRNAs were normalized to median expression of glyceraldehyde 3-phosphate dehydrogenase (GAPDH).

10

Data in Figure 1 indicate that expression levels for both IFN- $\lambda$ 2 and IFN- $\lambda$ 3 are higher in carriers of a genotype having the HR allele for rs8099917 ( $P < 0.04$ ). The genotype may alter expression in different contexts and with different stimulation *e.g.*, as indicated in Table 1, such as by altering one or more of mRNA splicing, mRNA turnover, mRNA half-life, mRNA stability, affinity of the encoded cytokine for its cognate receptor. Any one or more of these factors may contribute to improved viral clearance for subject having the high response (HR) haplotype. In any event, the data indicate functional significance in the correlation between haplotype and therapeutic efficacy of pegylated interferon-alpha (IFN- $\alpha$ ) and ribavirin against HCV.

20

*Clinical relevance*

Current therapies for HCV-1 employing immunomodulatory compositions such as IFN and ribavirin can produce serious adverse reactions and, in any event, produce low virus clearance in about 50% of infected patients. Accordingly, there is a clear benefit to providing diagnostic and prognostic methods to identify those subjects that are less likely to respond to therapy, thereby avoiding their discomfort. Such diagnostics and prognostic methods also provide a basis for suggesting adjunct or alternative therapies to those patients that are less likely to respond to conventional therapy.

25

The low response haplotype identified in this study is carried by 70% of northern Europeans, clearly indicating the extent of the problem faced by the pharmaceutical industry for effective therapy of this disease alone

- 5 The definition of SNPs, and associations between specific allelic variants at the loci identified in this study, have clear clinical relevance to the diagnosis and treatment using immune response modulators such as interferons, ribavirin and combinations thereof. For example, the identification of a subject carrying a low response (LR) allele at a SNP position identified in this study indicates a reduced likelihood of clearing a
- 10 virus such as HCV compared to a subject that is a non-carrier of the same allele. Similarly, the identification of a subject carrying a high response (HR) allele at a SNP position identified in this study indicates an enhanced likelihood of clearing virus compared to carriers of the LR allele. For example, 82% of GG homozygotes at the rs8099917 locus fail to clear HCV, whereas only 44% of TT homozygotes at this locus
- 15 fail to clear HCV. Standard genotyping and haplotyping methods as described herein may be employed to determine the likelihood of a response to therapy in a subject. Thus, the data provided herein provide the means to identify those subjects, including 50% of Europeans, who may clear virus on therapy, and those who do not. Using the SNPs identified herein, including the HR haplotype and LR haplotype associations,
- 20 nearly 90% of subjects capable of having high response(s) to conventional therapy can be identified by their genotype.

The data presented in this study also suggest the broad applicability of a diagnostic/prognostic assay based on IFN- $\lambda$ 3 genotyping and/or haplotyping to the

25 context of virus infections other than HCV. First, the association of IFN- $\lambda$ 3 with viral clearance is consistent with functionality of IFN- $\lambda$ 3 as an antiviral protein, the responsiveness of IFN- $\lambda$ 3 expression to Type 1 interferons such as IFN- $\alpha$  and IFN- $\beta$  (Li *et al*, *J. Leukocyte Biol.*, online publication DOI:10.1189/jlb.1208761 (Apr 30, 2009), and the observation that expression of IFN- $\lambda$ 3 is up-regulated in hepatocytes

30 and PBMCs of HCV-infected patients (Mihm *et al*, *C. Lab. Invest.* 84, 1148-1159, 2004). Second, IFN- $\lambda$ 3 is up-regulated by viral infection and by other interferons in

hepatocytes and other cells *e.g.*, Sirén *et al.*, *J. Immunol.* 174, 1932-1937 (2005), Ank  
*et al.*, *J. Virol.* 80, 4501-4509 (2006), and Doyle *et al.*, *Hepatology* 44, 896-906 (2006),  
and protects against HCV in an *in vitro* system *e.g.*, Robek *et al.*, *J. Virol.* 79, 3851-  
3854 (2005) and Marcello *et al.*, *Gastroenterology* 131, 1887-1898 (2006), as well as  
5 other RNA viruses *in vivo e.g.*, Ank *et al.*, *J. Virol.* 80, 4501-4509 (2006) and Ank *et al.*,  
*J. Immunol.* 180, 2474-2485 (2008). IFN- $\lambda$ 3 also regulates similar genes to IFN- $\alpha$   
via JAK/STAT signalling, however is more specific in its tissue targets. Proceeding on  
this basis, it is reasonable to conclude that IFN- $\lambda$ 3 provides the basis for diagnosis for  
those medical indications currently treated using IFNs. It is also reasonable to  
10 conclude that IFN- $\lambda$ 3 provides the basis for alternative therapies to those employing  
other IFNs such as IFN- $\alpha$  or IFN- $\lambda$ 1, *e.g.*, for those medical indications compatible  
with IFN- $\lambda$ 3 expression and activity.

Other associations described herein that are not linked to the IFN- $\lambda$  cluster are also  
15 strong indicators of virus clearance, as supported by the available data. The  
associations with SNPs linked to IL-21R on chromosome 16, caspase-1 (CASP-1) on  
chromosome 11 and an HLA pseudogene cluster on chromosome 6 are particularly  
interesting. For example, IL-21 promotes T cell proliferation and viral clearance, and  
is structurally similar to IFN- $\lambda$  in terms of exon structure, wherein the alpha helices are  
20 encoded by separate exons. Additionally, CASP1 activates IL-1 which then promotes  
the inflammatory cascade, and inhibits HCV replication *in vitro e.g.*, Zhu *et al.*, *J.*  
*Virol.* 77, 5493-5498 (2003). Proceeding on this basis, it is reasonable to conclude that  
the other associations described herein provide the basis for diagnosis/prognostic  
assays in the context of any medical indications currently treated using  
25 immunomodulatory compositions other than IFNs and/or ribavirin *e.g.*, compositions  
comprising IL-1.

### Example 2

Associations of HLA-C alleles with response to therapy of chronic hepatitis C using an immunomodulatory composition comprising an interferon (IFN)

- 5 A cohort of about 300 *e.g.*, 301, Australian Caucasians infected with HCV were genotyped for HLA-C alleles and associations between specific HLA-C alleles with response to therapy comprising interferon and ribavirin were determined using statistical procedures described herein.
- 10 Data are presented in Figure 2 and in Tables 8-10. The data presented in Table 8 show the associations between specific HLA-C1 alleles and specific HLA-C2 alleles and response to therapy. The data presented in Table 9 and Table 10 show that subjects who are homozygous for HLA-C2 alleles exhibit a low response to therapy compared to subjects who carry at least one HLA-C1 allele. These data indicate that HLA-C2
- 15 homozygosity is associated with a low response to therapy.

### Example 3

- Interactions *e.g.*, additive interaction or epistatic interaction between IL28B alleles and
- 20 HLA-C alleles in response to therapy of chronic hepatitis C using an immunomodulatory composition comprising an interferon (IFN)

#### *Combined effects of HLA-C and rs8099917*

- The cohort of subjects referred to in Example 2 were also genotyped for the IL28B
- 25 rs8099917 polymorphism, essentially as described in Example 1, to determine whether or not there is any interaction *e.g.*, additive interaction or epistatic interaction, between HLA-C genotype and rs8099917 genotype in determining the response to antiviral therapy.

- 30 Data presented in Figure 2 and Table 11 indicate that HLA-C2 homozygous individuals who are also carriers of the LR allele *i.e.*, G allele, linked to rs8099917, have only

about 20% likelihood of responding to antiviral therapy ( $p = 2 \times 10^{-4}$ , OR 7.3), and that subjects who are homozygous for HLA-C2 and homozygous for the LR allele of rs8099917 are even less likely to respond to therapy ( $p = 4.86 \times 10^{-5}$ ). In contrast, subjects who are carriers for at least one HLA-C1 allele and are homozygous for the  
5 HR allele of rs8099917 have about 60% likelihood of responding to therapy.

To establish reproducibility in this association, a second cohort of 483 Australian Caucasians infected with HCV were genotyped for HLA-C and rs8099917 alleles and associations between specific HLA-C alleles with response to therapy and interactions  
10 *e.g.*, additive interaction or epistatic interactions, between rs8099917 and HLA-C2 in modulating the response to therapy were determined as before. Data presented in Table 12 and Figures 3 and 4 confirm the associations between HLA-C2 alleles and response to therapy for the second cohort and the combined cohorts, and strengthen the conclusion that subjects who are homozygous for HLA-C2 alleles exhibit a low  
15 response to therapy compared to subjects who carry at least one HLA-C1 allele. These data also confirm the interactions *e.g.*, additive interaction or epistatic interaction between the rs8099917 polymorphism and HLA-C2 in response to antiviral therapy.

Data presented in Table 13 hereof also indicate that subjects who chronically-infected  
20 with HCV and being treated with IFN- $\alpha$  and ribavirin, have about 82.2% likelihood of not responding to treatment with a sustained viral response if they are homozygous for a HLA-C2 genotype and carry at least one LR allele linked to rs8099917. However, the predictive value of this test may not be improved substantially by further stratifying the population by testing for homozygosity of the LR allele of rs8099917. In contrast  
25 to the high positive predictive value of HLA-C2-homozygosity combined with rs8099917 LR genotyping, the presence of at least one HLA-C1 allele in combination with at least one LR allele at rs8099917 does not have such a high predictive value. This most likely reflects an interaction between HLA-C2 and the LR allele of rs8099917.

Comparison of the data in Tables 7 and 13 also suggests that the combination of HLA-C2 with one LR allele at rs8099917 provides for greater population coverage (sensitivity) than testing for homozygosity of the corresponding LR allele without significant compromise to positive predictive value or assay specificity. For example, only up to about 6% of the population are homozygous for the LR allele of rs8099917, whereas 16.5% of the population are homozygous for HLA-C2 and carry at least one LR allele of rs8099917, however both tests provide about 80-82% positive predictive value and about 96-98% specificity. Thus, there is a real advantage in combining HLA-C2 and rs8099917 LR genotyping in determining treatment outcome.

10

*Combined effects of HLA-C and rs4803221*

The cohort of subjects referred to in Example 2 were also genotyped for the IL28B rs4803221 polymorphism and HLA-C alleles, to determine whether or not there is any interaction e.g., additive interaction or epistatic interaction between the IL28B marker and HLA-C1 and/or HLA-C2 in response to antiviral therapy.

Data presented in Table 13 hereof also indicate that subjects who chronically-infected with HCV and being treated with IFN- $\alpha$  and ribavirin, have about 83.7% likelihood of not responding to treatment with a sustained viral response if they are homozygous for a HLA-C2 genotyping and carry at least one LR allele linked to rs4803221. This is better than the predictive value of at least one LR allele of rs8099917 combined with homozygosity of HLA-C2 alleles. Similarly, tests for homozygosity of LR alleles at rs4803221 have greatly-improved prognostic accuracy when combined with HLA-C2 homozygosity than do tests for homozygosity of LR alleles at rs8099917 when combined with HLA-C2. Notwithstanding the low frequency of HLA-C2/rs4803221-LR double-homozygotes in the population, the predictive value of this test also appears to be improved substantially by further stratifying the population by testing for homozygosity of the LR allele of rs4803221. In contrast to the high positive predictive value of HLA-C2 homozygosity combined with rs4803221 LR genotyping, the presence of at least one HLA-C1 allele in combination with at least one LR allele at

rs4803221 does not have a high predictive value. This most likely reflects an interaction between HLA-C2 and the LR allele of rs4803221.

Comparison of the data in Tables 7 and 13 also suggests that the combination of HLA-C2 with one LR allele at rs4803221 provides for greater population coverage (sensitivity) than testing for homozygosity of the corresponding LR allele without significant compromise to positive predictive value or assay specificity. For example, only up to about 8% of the population are homozygous for the LR allele of rs4803221, whereas 16.5% of the population are homozygous for HLA-C2 and carry at least one LR allele of rs4803221, however both tests provide about 84-88% positive predictive value and about 96-99% specificity. Thus, there is a real advantage in combining HLA-C2 and rs4803221 LR genotyping in determining treatment outcome.

Data presented in Table 13 also suggest that a combined test for homozygosity of HLA-C2 alleles and at least one LR allele at rs4803221 and at least one LR allele at rs8099917 may provide population coverage of about 28%, a specificity of more than 96% and a positive predictive value of more than 82%.

*Combined effects of HLA-C and rs12979860*

The cohort of subjects referred to in Example 2 were also genotyped for the IL28B rs12979860 polymorphism and HLA-C alleles, to determine whether or not there is any interaction e.g., additive interaction or epistatic interaction between the IL28B marker and HLA-C1 and/or HLA-C2 in response to antiviral therapy.

Data presented in Table 13 hereof also indicate that subjects who chronically-infected with HCV and being treated with IFN- $\alpha$  and ribavirin, have about 78% likelihood of not responding to treatment with a sustained viral response if they are homozygous for a HLA-C2 genotype and carry at least one LR allele linked to rs12979860. Notwithstanding that this performance may not be as high as that observed for rs8099917 or rs4803221, the minor (LR) allele is more abundant than for the other markers, and a wider population coverage may offset the disadvantage of a reduced

predictive value. Moreover, tests for homozygosity of LR alleles at rs12979860 have greatly-improved prognostic accuracy when combined with HLA-C2 homozygosity than do tests for homozygosity of LR alleles at rs8099917 when combined with HLA-C2. Notwithstanding the low frequency of HLA-C2/rs12979860-LR double-  
5 homozygotes in the population, the predictive value of this test also appears to be improved substantially by further stratifying the population by testing for homozygosity of the LR allele of rs12979860. In contrast to the high positive predictive value of HLA-C2 homozygosity combined with rs12979860 LR genotyping, the presence of at least one HLA-C1 allele in combination with at least one LR allele at  
10 rs12979860 does not have a high predictive value. This most likely reflects an interaction between HLA-C2 and the LR allele of rs12979860.

Comparison of the data in Tables 7 and 13 also suggests that the combination of HLA-C2 with one LR allele at rs12979860 provides for enhanced specificity and predictive  
15 value than testing for homozygosity of the corresponding LR allele.

Data presented in Table 13 also suggest that a combined test for homozygosity of HLA-C2 alleles and at least one LR allele at rs12979860 and at least one LR allele at rs8099917 may provide population coverage of about 33%, a specificity of about 95%  
20 and a positive predictive value of about 80%. By additionally combining such a test with at least one LR allele of rs4803221, the population coverage may near 50%, with still higher levels of specificity and accuracy.

#### *Combined effects of HLA-C and rs7248668*

25 The cohort of subjects referred to in Example 2 were also genotyped for the IL28B rs7248668 polymorphism and HLA-C alleles, to determine whether or not there is any interaction e.g., additive interaction or epistatic interaction between rs7248668 and either HLA-C1 alleles or HLA-C2 alleles in response to antiviral therapy.

30 Data presented in Table 13 hereof also indicate that subjects who chronically-infected with HCV and being treated with IFN- $\alpha$  and ribavirin, have about 85% likelihood of

not responding to treatment with a sustained viral response if they are homozygous for a HLA-C2 genotype and carry at least one LR allele linked to rs7248668. This is better than the predictive value of at least one LR allele of rs8099917 combined with homozygosity of HLA-C2 alleles. Similarly, tests for homozygosity of LR alleles at  
5 rs7248668 have greatly-improved prognostic accuracy when combined with HLA-C2 homozygosity than do tests for homozygosity of LR alleles at rs8099917 when combined with HLA-C2. Thus, notwithstanding the low frequency of HLA-C2/rs7248668-LR double-homozygotes in the population, the predictive value of this test also appears to be improved substantially by further stratifying the population by  
10 testing for homozygosity of the LR allele of rs7248668. In contrast to the high positive predictive value of HLA-C2 homozygosity combined with rs7248668 LR genotyping, the presence of at least one HLA-C1 allele in combination with at least one LR allele at rs7248668 does not have a high predictive value. This most likely reflects an interaction between HLA-C2 and the LR allele of rs7248668.

15

Comparison of the data in Tables 7 and 13 also suggests that the combination of HLA-C2 with one LR allele at rs7248668 provides for greater population coverage (sensitivity) than testing for homozygosity of the corresponding LR allele without substantial compromise to positive predictive value or assay specificity. For example,  
20 only up to about 8% of the population are homozygous for the LR allele of rs7248668, whereas about 16.5% of the population are homozygous for HLA-C2 and carry at least one LR allele of rs7248668, however both tests provide about 84-88% positive predictive value and about 96-99% specificity. Thus, there is a real advantage in combining HLA-C2 and rs7248668 LR genotyping in determining treatment outcome.

25

Data presented in Table 13 also suggest that a combined test for homozygosity of HLA-C2 alleles and at least one LR allele at rs7248668 and at least one LR allele at rs4803221 may provide population coverage of about 28%, a specificity of more than 96% and a positive predictive value of more than 84%, and that the combination of  
30 homozygosity of HLA-C2 alleles and at least one LR allele at rs7248668 and at least one LR allele at rs8099917 may provide similar population coverage and specificity

albeit lower positive predictive value. A combined test for homozygosity of HLA-C2 alleles and at least one LR allele at rs7248668 and at least one LR allele at rs4803221 and at least one LR allele at rs8099917 may also provide population coverage of about 43%, a specificity of more than 96% and a positive predictive value of more than 83%.

- 5 A combined test for homozygosity of HLA-C2 alleles and at least one LR allele at rs7248668 and at least one LR allele at rs4803221 and at least one LR allele at rs8099917 and at least one LR allele at rs12979860 may also provide population coverage of about 61.6%, a specificity of more than 95% and a positive predictive value of more than 80% or 81% or 82%. By additionally combining such a test with at  
10 least one additional LR allele shown in Table 1 hereof, especially an LR allele in L.D. with rs8099917, the population coverage may be even higher, and still higher levels of specificity and accuracy may be achieved.

#### *Summary*

- 15 In particular, the data support the conclusions that:

1. HLA-C2 homozygotes are less likely to respond to pegylated interferon/ribavirin therapy than those with HLA-C1 genotypes (Odds Ratio 1.57);
2. Carriers of one or more LR alleles of IL28B SNPs are less likely to respond to  
20 pegylated interferon/ribavirin therapy than the corresponding HR genotypes, including the corresponding IL28B-HR homozygotes;
3. HLA-C2 homozygotes who are also carriers of one or more LR alleles of IL28B SNPs are much less likely to respond to drug therapy than other genotype combinations *e.g.*, see Figs 3 and 4, Tables 7 and 13; and
- 25 4. The positive predictive values of exemplary LR genotypes when combined with HLA-C2 homozygosity may be ranked as follows:

rs7248668=rs4803221>rs8099917>rs12979860.

## Example 4

Identification of improved IL28B SNPs and haplotypes for prediction of drug response in treatment of hepatitis C using massively parallel sequencing in a cross-sectional European cohort

5

Patient cohorts

All patients included in this study had been infected with genotype 1 HCV (the diagnosis of which was based on serology and serum viral RNA tests), had received a standard course of pegylated interferon and ribavirin (PegIFN/R), and had their post-treatment virological response determined six months after completion of therapy. All sustained virological responders (SVRs) and non-SVR cases had received therapy for 48 weeks except where HCV RNA was present with a  $<2$  log drop in RNA level after 12 weeks therapy. Patients were classified as having a sustained virological response (SVR) if they were determined as being HCV PCR negative, six months after completion of therapy. Patients were excluded from this study if they had been co-infected with either hepatitis B virus (HBV) or human immunodeficiency virus (HIV) or if they were not of European descent. Characteristics of each cohort are presented in Table 14.

20 Ethics statement and study subjects

Ethical approval was obtained from the Human Research Ethics Committees of Sydney West Area Health Service and the University of Sydney. All other sites had ethical approval from their respective ethics committees. Written informed consent was obtained from all participants.

25

Massively parallel sequencing

The study design is illustrated in Figure 5. DNA samples obtained from 100 of the 131 patients that responded to treatment with PegIFN/R *i.e.*, responders, were pooled to create a responder DNA pool, while DNA samples obtained from 99 of the 162 patients having a poor response or no response to treatment with PegIFN/R *i.e.*, non-responders, were pooled to create a non-responder DNA pool. DNA samples were not barcoded.

Long-range PCR was then used to amplify a continuous 100 kbp region of DNA containing the IL28A, IL28B, and IL29 genes. Specifically, 23 overlapping amplicons of sizes between 4800-5273 bp were used to amplify chr19:39,709,944-39,809,945 (human genome version hg19). The PCR products produced for each pool were then  
5 sequenced using an Illumina Genome Analyzer II flowcell, with 75 bp paired-end reads generated. Raw read quality was examined using FASTQC version 0.7.0 *e.g.*, as described by Andrews S: FASTQC. *A quality control tool for high throughput sequence data*, accessible at <http://www.bioinformatics.bbsrc.ac.uk/projects/fastqc/>. Genotype data generated has subsequently been deposited at the European  
10 Genomephenome Archive (EGA), accessible at <http://www.ebi.ac.uk/ega/>, which is hosted by the EBI, under Accession number [EBI:EGAS00001000096].

#### Sequence alignment and post-processing

Reads from each pool were aligned separately to hg19 using version 0.5.8a of Burrows-  
15 Wheeler Alignment (BWA) *e.g.*, as described by Li *et al.*, *Bioinformatics*, 25, 1754-1760 (2009). Up to four mismatches were permitted across the length of each read, including up to four mismatches in the 32 bp seed. Bases with Phred quality scores of less than three were trimmed from the ends of reads.

20 Reads corresponding to the target region were extracted using SAMtools *e.g.*, as described by Li *et al.*, *Bioinformatics*, 25, 2078-2079 (2009). Sequence reads with a mapping quality of zero *e.g.*, singletons that do not align to unique location in the genome or read pairs where neither end maps to a unique location, were discarded.

25 In view of the small size of the target region, the large number of haplotypes present, and the depth of sequencing, a substantial proportion of the read pairs were expected to align to the same location(s). However, as such duplicates were determined to be biological in nature, rather than mere PCR duplicates, removal of duplicates was not performed.

#### Variant calling and association testing using pooled DNA

Variant calling was performed using version 3 of CRISP *e.g.*, as described by Bansal V: *Bioinformatics*, 26, i318-324 (2010), employing a variant detection algorithm designed specifically for pooled DNA samples. Default settings were used (minimum  
5 read mapping/base quality to consider a read/base for variant calling 10;  $\geq 1$  read supporting variant must have mapping quality  $\geq 20$ ;  $\geq 4$  reads per pool; contingency table threshold  $p=10^{-3}$ ; quality values based  $p$ -value threshold  $10^{-5}$ ).

Fisher's exact test was employed to compare the proportion of each allele in responders  
10 and non-responders at variant sites. If the total number of allele counts for a cohort exceeded the number of chromosomes, allele counts were rescaled so that they summed to the number of chromosomes in the cohort.

#### Validation with GWAS SNP data

15 Genotype data for 15 SNPs within the 100 kbp target region previously generated as part of a GWAS study were compared to data estimated from pooled MPS data to compare the minor allele frequencies (MAFs), odds ratios and Fisher exact  $p$ -values. Correlation between the two data sets was then summarised using Spearman's rank correlation coefficient.

20

We also compared MAFs from pooled MPS to MAFs from CEPH genotypes, obtained from Utah residents with ancestry from northern and western Europe (CEU), for the 35 SNPs in the target region which were studied as part of the HapMap project *e.g.*, as described by The International HapMap Consortium, *Nature*, 449, 851-861 (2007).

25

#### Validation using individual genotyping

19 putative single nucleotide variants (SNVs) were chosen for validation using individual genotyping based on the MPS results and/or biological grounds. MPS SNPs were ranked by Fisher's exact test. SNPs with  $p$ -values exceeding  $10^{-3}$  were rejected,  
30 and only those SNPs supported by reads aligning in both directions and covered by at least 1000 reads in at least one pool were retained. Those SNPs failing Sequenom

genotyping algorithms *i.e.*, which excludes SNPs which are in highly homologous regions, have multiple genomic targets, or are close to other variants which may confound the genotyping, were also rejected. 19 SNPs were retained for validation, all included in dbSNP132. Genotype data for ten of those SNPs were available from  
5 previous studies, and genotype data for the remaining nine SNPs was generated in a single multiplex Sequenom reaction. The validation cohort comprised 905 samples obtained from six different cohorts from Australia, the UK, Germany and Italy. Three SNPs (rs8105790, rs8103142 and rs628973) were not genotyped for 324 samples and 43 samples were not genotyped for five SNPs (rs10853727, rs8109886, rs10853728,  
10 rs12980602, rs4803224).

Genotype data cleaning was performed using PLINK *e.g.*, as described by Purcell *et al.*, *American Journal of Human Genetics*, 81, 559-575 (2007). To account for obligate missingness, samples with a genotyping rate of below 89.47% (corresponding  
15 to less than 17/19 SNPs being assigned a genotype), and SNPs with (i) a genotyping rate below 90%, (ii) a minor allele frequency (MAF) below 1%, or (iii) a combined sample Hardy Weinberg test p-value below  $10^{-10}$ , were discarded. A Pearson's chi-square test was used to test for differential missingness of SNPs between responders and non-responders.

20

Single marker association analyses were also performed using PLINK. Genotypes from the six cohorts were combine and tests of association under allelic, genotypic, recessive and dominant genetic models were performed. A fixed-effects meta-analysis for the allelic test was also performed, although this was not deemed advisable under  
25 other genetic models as low or zero genotype counts resulted for some SNP/cohort combinations. Linkage disequilibrium (LD) was estimated and haplotype frequencies inferred using Haploview *e.g.*, as described by Barrett *et al.*, *Bioinformatics*, 21, 263-265 (2005). The frequencies of common haplotypes ( $\geq 1\%$ ) were compared between non-responders and responders using odds ratios, with 95% confidence intervals  
30 calculated using Woolf's formula and p-values calculated using Pearson's chi-square test.

## Results

### Massively parallel sequencing

From the test cohort, the responder pool generated 34,026,026 reads (17,013,013 pairs) whereas the non-responder pool generated 34,236,380 reads (17,118,190 pairs). Analysis of the raw sequence reads using FASTQC indicated overrepresentation of certain sequences, which were determined as corresponding to the ends of amplicons, *i.e.* primers and adjacent sequence.

### 10 Alignment

33,761,231 responder reads and 33,965,033 non-responder reads aligned back to hg19 (99.2% for both pools), of which 31,960,774 and 32,008,412 reads aligned back to the target region, respectively. Of these, 31,082,806 and 31,143,590 reads had unique alignments to the target region (not shown here).

15

Median coverage across the target region was determined as 19,200 for responders (range 0-171,197) and 20,220 for non-responders (0-177,376). 99.28% and 99.10% of targeted bases had coverage  $\geq 4$  for R and NR, respectively. Coverage was determined to be relatively low for one of the two pools for two of the 23 amplicons. Very low coverage was observed for the responders for the twentieth amplicon (chr19:39,789,495 to 39,794,767), whereas very low coverage was observed for the non-responders for the first amplicon (chr19:39,709,955 to 39,714,897). Inspection of the relevant PCR products showed only faint bands. Thus 2/46 amplicon pools were likely to have suffered PCR failures leading to low read counts. Fortunately, these amplicons were not located close to IL28B.

### Variant calling and association testing

CRISP detected 1221 SNVs and 118 indels (1339 variants in total).

#### Validation with GWAS SNP data and HapMap data

CRISP called all 15 of the SNPs individually typed as part of the GWAS as variants. The Spearman correlation coefficients of MAF estimates between the pooled MPS and individual genotyped data for the 15 SNPs was determined to be 0.99 (R) and 0.95 (NR). Odds ratios were determined to be well correlated ( $\rho=0.98$ ) while p-values were shown to be poorly correlated ( $\rho=0.35$ ).

31 of the 35 HapMap SNPs in the target region were called as variants; CEU MAFs of 0.013 or higher. The four HapMap SNPs not called as variants had CEU MAFs of 0.004, 0.005, 0.009 and 0.027 respectively. Spearman correlation coefficients of CEU MAFs with MAFs estimated from pooled MPS data were 0.84 (R) and 0.80 (NR).

#### Individual genotyping results

86 samples with a genotyping rate below 89.47% were discarded, leaving 819 samples remaining. All 19 SNPs genotyped in this study had genotyping rates above 90%, with 17 SNPs having genotyping rates of 99% or higher. SNP rs628973 was discarded due to a suspicious distribution of genotypes (98.25% of genotypes heterozygous, Hardy-Weinberg test for combined sample  $p=9.78 \times 10^{-136}$ ). The remaining 18 SNPs had Hardy-Weinberg test p-values ranging from  $3.09 \times 10^{-8}$  to 0.78, minor allele frequencies ranging from 0.036 to 0.47, and differential missingness p-values ranging from 0.095 to 1. Accordingly, association analyses using 819 samples with genotypes for up to 18 SNPs were performed. A meta-analysis of allelic test results determined that 12 of the SNPs were strongly ( $p<0.001$ ) associated with response..

25 The four SNPs that best tagged the IL28B region haplotypes were investigated further for their ability to predict response to therapy.

The association testing results using Fisher's exact test based on the BWA alignment and CRISP variant calls from the MPS data with rescaled read counts demonstrate a strong clustering of association signals around IL28B. The two problem amplicons where PCR failures were suspected, as previously discussed, were retained in the

analysis. Figure 6 shows that both of these regions to which these amplicons align produced spurious association signals with amplicon 1 producing the most significant association signals overall.

- 5 Association testing of the nineteen SNPs selected for individual genotyping in the independent combined cohort validate the MPS association testing results for the IL28B region, as well as previous GWAS results and other more recent re-sequencing results which are not MPS derived. Such consensus clearly implicate IL28B as the most likely causal gene. SNPs rs4803221 and rs7248668 were determined as having
- 10 the highest association for a recessive model with an OR=4.74 and 4.85 respectively (Figure 7). Surrounding SNPs also show clear recessive effects. These results may be explained by the haplotype effect as shown in Table 15.

#### Haplotypes and Linkage Disequilibrium

- 15 From the total cohort, six haplotypes were identified using Haploview (Table 15, Figure 8). Haplotype 2 showed the most significant association with failure to respond to therapy. This was tagged by rs8099917 and rs7248668, and largely by rs4803221. These SNPs also returned the highest odds ratios for homozygotes.

#### Predictive value for failure to respond.

- To evaluate the utility of the SNPs to predict treatment failure, the sensitivity, specificity, positive (for treatment failure, PPV) and negative predictive values for the SNP minor allele homozygotes were compared. The best four SNP minor allele homozygotes are shown in Table 16. The PPV for treatment failure is likely to be the
- 25 most useful parameter clinically. The SNPs rs4803221 GG (PPV of 77.1, CI 62.7-88.0) and rs7248668 AA (PPV 78.3, CI 63.6-89.1) both performed better than the SNPs currently used in testing rs8099917 GG (PPV 73.3, CI 58.1-85.4) and rs12979860 TT (PPV 68.3; CI 59.2-76.5) in this sample. However, confidence intervals indicate that they may perform better, worse or equivalently at a population level.

As demonstrated in Example 3, a combination of carrier status for the rs8099917 minor allele with HLA-C C2C2 homozygosity predicts treatment failure (PPV of approx. 80%) better than either genotype alone, and more people have this genotype than are homozygous for the minor alleles. In the present study, all haplotype 2 SNPs  
5 (rs4803221 G, rs7248668 A, and rs8099917 G) were determined to perform similarly (PPV 78.6, 79.7, 80.3), and better than rs12979860 T (PPV 73.1). The maximum proportion of the population identified as having a non-responder genotype is determined by including those with minor allele homozygosity (rs4803221 GG, PPV 77.1) with those who are carriers of rs4803221 G and are HLA-C C2 homozygotes  
10 (PPV 78.6). In the present case, this was determined to be 7.7% of the responder cohort, and 19.9% of the non-responders. Although SNP rs1297860 T homozygosity plus T carriers with HLA-C C2 homozygosity was determined to predict a higher proportion of people who will fail to respond to treatment (33.2% of non-responders), this combination was determined to mis-classify more responders (17.1% responders).

15

*In silico* analysis of transcription factor binding sites and methylation sites in the proximal promoter region

A CpG island was previously identified on the UCSC human genome draft by Miliak and Hillier (Figure 9), encompassing SNPs rs12978960 and rs4803221. The major  
20 allele for each SNP comprises the C of a CpG dinucleotide. Transcription factors that might differentiate between the haplotypes were sought using the program Ali Baba (Figure 9), which identifies transcription factor sites based on core recognition motifs. Several transcription factors that might differentiate between the haplotypes were identified, and these varied according to SNP present on the haplotype.

25

**Discussion**

Data presented herein identified eighteen SNPs in the putative promoter region of IL28B using massively parallel sequencing on pooled responders and non-responders, which predicted response to PegIFN/R. These eighteen SNPs collapse to six  
30 haplotypes that predict response according to tagging by their minor alleles. The minor allele for SNPs rs8099917, rs7248668, and rs4803221 were all found on haplotype 2,

which had the highest OR (recessive model) for predicting treatment failure. The minor allele of SNP rs4803221 may also be found on the rare haplotype 6. Minor alleles from SNPs rs12979860, rs11881222, rs688187, rs12982533, rs35790907, rs12982533, rs12980275 were also found on haplotype six and two to three others  
5 (with MAF>1%).

The LD block varies between ethnic groups. Japanese and Chinese have fewer common haplotypes, with virtually complete linkage between rs8099917, rs12978960, and the 3' end of the gene. In the population of the instant study, any of the SNPs  
10 predict treatment failure with similar precision. People of European descent, Hispanics and Indians have more common haplotypes. In the former two it has been determined that haplotype 2 best predicts failure to respond, with homozygotes for haplotype 2 best predicting treatment failure. There is some recombination between the SNPs tagging haplotype 2, and then the best two are rs7248668 (next to rs8099917) and rs4803221  
15 (next to rs12979860, also on the rare hapotype 6). The African population has the shortest LD block, with a boundary between rs12979860 and rs8099917, and previous investigation has shown that the major effect of this African haplotype lies between this point and the 3' end of the IL28B *e.g.*, Ge *et al.*, *Nature*, 461, 399-401 (2009). The causative SNPs or SNP, or other genetic variant, is therefore most likely to be in this  
20 section of the gene, which includes promoter variants, intronic, codon changing, and 3'untranslated SNPs, any of which could be, or could contribute with others, to the functional effect of the haplotype. However, Ge *et al* (2009) identified an effect of rs8099917 independent of this block in African Americans, indicating multiple, variable, haplotype effects on gene function.

25

#### The Causative Haplotype

Each of the haplotypes were determined to have putative immune transcription factor sites distinguishing them from one another. Notably, the haplotype 2 SNP, rs4803221, resides in the CpG island, and removes a CpG site. The SNP rs12978960 is the only  
30 other common SNP also in this region, and the variant on haplotype 2 also removes a CpG site. Therefore, two potential methylation sites are missing from haplotype 2, and

none or one methylation site from the haplotypes corresponding to response. Methylated DNA is known to be resistant to unfolding and, as such, corresponds to reduced expression. Proceeding on this basis, it is reasonable to hypothesize that increased methylation may lead to reduced expression of IL28B, and the interferon sensitive genes (ISGs) upregulated by it, in the responder haplotype; which is then responsive to interferon alpha stimulation on therapy. Two studies have identified that IL28B non-responders have high ISG expression in infected hepatocytes, and that high ISG levels independently predict poor response to therapy *e.g.*, as described by Honda *et al.*, *Gastroenterology*, 139, 499-509 (2010). However the present inventors, and others, have found that the non-responder haplotype is associated with reduced expression in peripheral blood of healthy controls *e.g.*, Suppiah *et al.*, *Nature Genetics* 41, 1100-1104 (2009) and Tanaka *et al.*, *Nature Genetics*, 41, 1105-1109 (2009). On the other hand, other authors have also found no evidence for an effect of IL28B genotype on ISG expression in uninfected hepatocytes *e.g.*, O'brian T.R., *Nature Genetics*, 41, 1048-1050 (2009), suggesting that expression of IL28B and ISGs may be context dependent.

Based on the data present herein, haplotype 2 is clearly the major causative haplotype for predicting response to therapy. The causative SNP or SNPs will probably best be sought in the African American population, where the linkage disequilibrium is lowest, allowing breaking up of the large haplotype block and thus allowing inference of the relative effects of each SNP. The task of demonstrating the basis for causality, even of haplotypes, is not trivial, and thus far has not been achieved for most of the genetic associations currently identified by GWAS. In the short term, diagnostic tests are likely to be dependent on haplotype tagging SNPs, in combination with other parameters such as IP10, 25 hydroxy vitamin D3, and HLA-C genotype. From genotyping, SNP rs4803221 and HLA-C C2C2 provides the best predictive value.

#### Clinical relevance

Therapy for HCV infection is currently undergoing a radical transformation with the advent of Direct Acting Antivirals (DAAs) such as boceprvir and telaprvir to the

PegIFN/R backbone. Cure rates are substantially higher (20-30%), and these drugs have the potential to reduce treatment duration. However, enthusiasm for these regimens is tempered by the substantially lower cure rates (~30%) in previous PegIFN/R null responders. In all those failing single DAA-based therapy, future treatment with multiple DAA-based combinations with or without PegIFN/R may be compromised by the development of drug resistance. Furthermore, HCV eradication using single DAA-based strategies, particularly in previous treatment failures, appears to be IL28B genotype dependent. In this context, predicting non-response rather than success is paramount since the former should perhaps have therapy deferred until multiple DAA-based combinations become available. The present results therefore provide a strong rationale for the use of rs4803221 in combination with the HLA-C genotype, such that those with non-response genotypes be considered for future regimens rather than single DAA-based therapy, signalling the emergence of truly personalized treatment for HCV infection.

15

#### Resequencing for GWAS with MPS

Due to cost constraints the current approaches for GWAS, re-sequencing usually employ pooling strategies. Large pool sizes makes it difficult to differentiate between rare variants and sequence errors, while the variable representation of each haplotype among those sequenced makes MAF estimates more variable and complicates association testing. However, if analysis is performed carefully it is possible to not only identify variants, but to also meaningfully assess the association signals. Data presented herein represent the first time that this region of the genome has been re-sequenced with MPS to discover novel variants that could determine the response to hepatitis C treatment. No rare variants tagging haplotype 2 were identified, but this does not exclude the existence of a synthetic association via multiple rare variants. Such rare variants would have to be predominantly on haplotype 2.

MPS resequencing in GWAS is not without problems. Targeting of the respective region is required, usually via PCR, for smaller association signals. PCR can lead to amplification artefacts that can be alleviated, at least in part, with advanced lab

protocols which may result in more uniform coverage than what was achieved in the present study. Failure to assess amplicon pool performance may lead to false positive association signals in MPS resequencing studies. Genomic regions with high homology, such as the one that IL28B is located in, have low mapability, resulting in  
5 reduced ability to detect new variants. Another problem is that short reads can misalign to a different version of the homologous sequence than the one they were derived from, resulting in alignments with mismatches that are incorrectly identified as putative variants. Using an alignment program such as BWA, which allows for gapped alignment around indels, is crucial. Thus the present study has provided important  
10 insights for future resequencing studies following up GWAS hits.

Table 1

## Summary of SNPs associated with response to therapy

| SNP        | Chromosome | Position <sup>1</sup> | Location <sup>2</sup>         | SNP effect       | Sequence comprising SNP                                           | SEQ ID NO: |
|------------|------------|-----------------------|-------------------------------|------------------|-------------------------------------------------------------------|------------|
| rs4803224  | 19         | 44444854              | IL28A/IL28B intergenic region | expression level | aaaaaaaaaagagaataatcctggcag[C/G]tgggtggctgcagctccagctgcttag       | 1          |
| rs12980602 | 19         | 44444660              | IL28A/IL28B intergenic region | expression level | atattcataaacaatatgaagccagaga[C/T]agctcgtctgagacacagatgaacaaaac    | 2          |
| rs10853728 | 19         | 44436986              | IL28A/IL28B intergenic region | Weak             | tgtctgtaagcagcctggggagatgtgggc[C/G]taagcttggtaggagatgagagctgtctt  | 3          |
| rs8099917  | 19         | 44435005              | 5'-end of IL28B               | expression level | cctcctttgtttcttctctgtgagcaat[G/T]tcacccaaattggaaccatgctgtatatacag | 4          |
|            |            |                       |                               | HR allele=T      | cctcctttgtttcttctctgtgagcaatTtcacccaaattggaaccatgctgtatatacag     | 5          |
|            |            |                       |                               | LR allele=G      | cctcctttgtttcttctctgtgagcaatGtcacccaaattggaaccatgctgtatatacag     | 6          |
| rs8113007  | 19         | 44434943              | 5'-end of IL28B               | expression level | ttaaagtaagctctgtatttcacctctgg[A/T]ggtaaatatttttaacaattgtcactgt    | 7          |
| rs8109889  | 19         | 44434610              | 5'-end of IL28B               | expression level | caatttccaacaagcatcctgcccagg[C/T]gctcgtcgtctcctcaatcctcttttg       | 8          |
| rs8109886  | 19         | 44434603              | 5'-end of IL28B               | expression level | ttctattcattttccaacaagcatcctg[A/C]cccaggctcgtcgtctcctcaatc         | 9          |
|            |            |                       |                               | HR allele=C      | ttctattcattttccaacaagcatcctgCcccaggctcgtcgtctcctcaatc             | 10         |
|            |            |                       |                               | LR allele=A      | ttctattcattttccaacaagcatcctgAccaggctcgtcgtcgtcctcaatc             | 11         |
| rs61599059 | 19         | 44434538              | 5'-end of IL28B               | expression level | gtctgtttcttcttctctctctctct[*]/CT]gttctgtctcgtctggtggactcca        | 12,13      |

1, Chromosome positions are derived from Hapmap project data release 27.

2, Gene locations were obtained by scanning  $\pm 100$  kb from the associated SNP

HR allele, Allele associated with higher response to therapy.

LR allele, Allele associated with lower response or no response to therapy.

Table 1 continued  
Summary of SNPs

| SNP        | Chromosome | Position <sup>1</sup> | Location <sup>2</sup> | SNP effect       | Sequence comprising SNP                                            | SEQ ID NO: |
|------------|------------|-----------------------|-----------------------|------------------|--------------------------------------------------------------------|------------|
| rs34567744 | 19         | 44434535              | 5'-end of IL28B       | expression level | tggtctggtttcttcttctctctc[*/CT]ctgttccgtctctgtctctggcgtgact         | 14,15      |
| rs10642510 | 19         | 44434534              | 5'-end of IL28B       | expression level | tgcttgcttcttcttctctctc[*/CT/TC]ctctgttctgtctctgtctctgggt           | 16,17,18   |
| rs10643535 | 19         | 44434531              | 5'-end of IL28B       | expression level | tctgtgtcttgcttcttcttctctctc[**/CT]ctctctgttctgtctctgtctctggcgtg    | 19,20      |
| rs34593676 | 19         | 44434523              | 5'-end of IL28B       | expression level | agcgtctctctctgtgtgtgtgtgtgtctt[**/TC]tctctctctctctgttctctgtctctgtc | 21,22      |
| rs25122122 | 19         | 44434521              | 5'-end of IL28B       | expression level | tcagcgtctctctctgtgtgtgtgtgtgtctt[*/T]tttctctctctctctgttctctgtctctg | 23,24      |
| rs35407108 | 19         | 44434307              | 5'-end of IL28B       | expression level | gccgtgggcaacaaagtgaactcgtctc[*/A]aaaaaaacacacacaaaggagggttc        | 25,26      |
| rs59211796 | 19         | 44434282              | 5'-end of IL28B       | expression level | gccgagatcacgcattgcactccagcgtg[A/G]gcaacaaagfgaaactcgtctcaaaaa      | 27         |
| rs62120529 | 19         | 44434043              | 5'-end of IL28B       | expression level | aaaaagacacaaaccaggcacagtcgctc[A/G]tgccctgaatccacagcacttgggagggccg  | 28         |
| rs62120528 | 19         | 44433258              | 5'-end of IL28B       | expression level | cttgaggtaggaggtcaataaccagcctga[A/C]caacatggcaaacacctgtctctactagaa  | 29         |
| rs12983038 | 19         | 44432964              | 5'-end of IL28B       | expression level | ggagaggaggttggagcccagggttc[A/G]agaccagcctgggcaataatagtgagaccct     | 30         |
| rs10853727 | 19         | 44432303              | 5'-end of IL28B       | weak             | tttgcgaacatacatatgaagaggca[C/T]gcttaigatcgcacctgcgtctggaggttg      | 31         |
| rs7254424  | 19         | 44432022              | 5'-end of IL28B       | expression level | aattctggattacagcagatcattgc[A/G]ccctggcctcattatttttaaacogtttt       | 32         |

1, Chromosome positions are derived from Hapmap project data release 27.

2, Gene locations were obtained by scanning  $\pm 100$  kb from the associated SNP

HR allele, Allele associated with higher response to therapy.

LR allele, Allele associated with lower response or no response to therapy.

Table 1 continued  
Summary of SNPs

| SNP        | Chromosome                                                                    | Position <sup>1</sup> | Location <sup>2</sup> | SNP effect       | Sequence comprising SNP                                            | SEQ ID NO: |
|------------|-------------------------------------------------------------------------------|-----------------------|-----------------------|------------------|--------------------------------------------------------------------|------------|
| rs1549928  | 19                                                                            | 44431549              | 5'-end of IL28B       | expression level | gaagcaagaagaagagaaacacacagtaga[A/G]acaggagacagacaattggaaaccgagt    | 33         |
| rs34347451 | 19                                                                            | 44431529              | 5'-end of IL28B       | expression level | gggatggctgcccctccacacacagggttcc[* /A]aaatgtctctgtcctgtttctactgtct  | 34,35      |
| rs35814928 | 19                                                                            | 44431477              | 5'-end of IL28B       | expression level | tcgggatcccatctcgggtgtgaggacttc[* /A]aacccagggttgccctgtgcccgggagagc | 36,37      |
| rs4803222  | 19                                                                            | 44431193              | 5'-end of IL28B       | expression level | ggcgtgaggacacacacacacagtgagg[C/G]agagagtgaggagccggccccctctcgcct    | 38         |
| rs11322783 | 19                                                                            | 44430995              | 5'-end of IL28B       | expression level | agtgcgagagcaggcagcgcggggggcc[* /T]cgcgatcacccgtgcacagggacccacagcc  | 39,40      |
| rs4803221  | 19                                                                            | 44430969              | 5'-end of IL28B       | expression level | cagcgtccggggctccagcggagcggtagtg[C/G]gagagcagcagcgcggggggccctctgc   | 41         |
|            |                                                                               |                       |                       | HR allele=C      | cagcgtccggggctccagcggagcggtagtgCgagagcagcagcgcggggggggccctctgc     | 162        |
|            |                                                                               |                       |                       | LR allele=G      | cagcgtccggggctccagcggagcggtagtgGgagagcagcagcgcggggggggccctctgc     | 163        |
| rs12979860 | 19                                                                            | 44430627              | 5'-end of IL28B       | expression level | tgtactgaaccaggaggagctcccgaggcg[C/T]gaaccagggttggaattgcaactcgcgctcc | 42         |
|            |                                                                               |                       |                       | HR allele=C      | tgtactgaaccaggaggagctcccgaggcgCgaaccagggttggaattgcaactcgcgctcc     | 164        |
|            |                                                                               |                       |                       | LR allele=T      | tgtactgaaccaggaggagctcccgaggcgTgaaccagggttggaattgcaactcgcgctcc     | 165        |
| rs12971396 | 19                                                                            | 44429706              | 5'-end of IL28B       | expression level | gaagaccacgctggcttggcgcaccagg[C/G]gagtcctggagccagggaggaggagggcagcg  | 43         |
| rs11672932 | 19                                                                            | 44429556              | 5'-end of IL28B       | expression level | tcgcggccagcccaatggagcacag[C/G]agctgtttcggcagccaatggcggtgg          | 44         |
| 1,         | Chromosome positions are derived from Hapmap project data release 27.         |                       |                       |                  |                                                                    |            |
| 2,         | Gene locations were obtained by scanning $\pm 100$ kb from the associated SNP |                       |                       |                  |                                                                    |            |
| HR allele, | Allele associated with higher response to therapy.                            |                       |                       |                  |                                                                    |            |
| LR allele, | Allele associated with lower response or no response to therapy.              |                       |                       |                  |                                                                    |            |

Table 1 continued  
Summary of SNPs

| SNP        | Chromosome | Position <sup>1</sup> | Location <sup>2</sup> | SNP effect       | Sequence comprising SNP                                      | SEQ ID NO: |
|------------|------------|-----------------------|-----------------------|------------------|--------------------------------------------------------------|------------|
| rs11882871 | 19         | 44429451              | 5'-end of IL28B       | expression level | tccctgtagaagagaccgctctctt[A/G]tatctgagacagtgatccaaagtcag     | 45         |
| rs56215543 | 19         | 44429428              | 5'-end of IL28B       | expression level | gataaagagagagcgggctctcttac[A/G]gggaaagagaccacagttccaggaa     | 46         |
| rs12979731 | 19         | 44429353              | 5'-end of IL28B       | expression level | tccagagctcaagtttttctctgccal[C/T]agcaaccgttgagggtcgtacaaag    | 47         |
| rs2020358  | 19         | 44428927              | 5'-end of IL28B       | expression level | cagaccaggagactcaggtggccgtgag[G/T]ttcagttc'gacctgccagttaatt   | 48         |
| rs34853289 | 19         | 44428781              | 5'-end of IL28B       | expression level | tcaataagaccatactaggaacctcag[C/T]tggagagtttaaacgtgactcaac     | 49         |
| rs8107030  | 19         | 44428559              | 5'-end of IL28B       | expression level | gggtgccgtcttctcttaggggaagttc[A/G]tggcagtggtgaagagcagtggtcttg | 50         |
| rs41537748 | 19         | 44428498              | 5'-end of IL28B       | expression level | aggctctgtcacaaga[C/T]tgagggtgacgaagg                         | 51         |
| rs59702201 | 19         | 44428148              | 5'-end of IL28B       | expression level | gcataataatataatataatataat[A/TAT]tttgagacagggtctgttctggctcac  | 52,53      |
| rs2596806  | 19         | 44428010              | 5'-end of IL28B       | expression level | taagacaggggtctcactcgtcactg[C/G]agtgcgaatggcagatcacacagctcac  | 54         |
| rs2569377  | 19         | 44427950              | 5'-end of IL28B       | expression level | gtaacctacaggaaggtatgttccca[A/G]gaggattccaccigtctgtgtttgt     | 55         |
| rs4803219  | 19         | 44427759              | 5'-end of IL28B       | expression level | ctgagctccatggggcagcttttattc[C/T]ctgacagaaggcagtcgccagctgat   | 56         |

1, Chromosome positions are derived from Hapmap project data release 27.

2, Gene locations were obtained by scanning  $\pm 100$  kb from the associated SNP

HR allele, Allele associated with higher response to therapy.

LR allele, Allele associated with lower response or no response to therapy.

Table 1 continued  
Summary of SNPs

| SNP        | Chromosome                                                                    | Position <sup>1</sup> | Location <sup>2</sup>    | SNP effect                                                            | Sequence comprising SNP                                       | SEQ ID NO: |
|------------|-------------------------------------------------------------------------------|-----------------------|--------------------------|-----------------------------------------------------------------------|---------------------------------------------------------------|------------|
| rs28416813 | 19                                                                            | 44427484              | 5'-end/intron 1 of IL28B | expression<br>level/mRNA<br>stability/turnover/<br>alternate splicing | cagagagaaaggagcctgagggaaag[C/G]agaggctgccccctgaggcaggggc      | 57         |
| rs630388   | 19                                                                            | 44427442              | exon 1 of IL28B          | silent mutation in<br>codon for Thr6 of<br>IL28B                      | agcaccagcactggcctgagtcagtcctcc[A/G]gtcctgctgtgtcaccagagagaaag | 58         |
| rs629976   | 19                                                                            | 44427365              | exon 1 of IL28B          | missense: R32H<br>mutation in IL28B                                   | tggagcagttctctgtgccaggtctc[A/G]cggggctctccggatgcaaggggct      | 59         |
|            |                                                                               |                       |                          | IL28B-His32 allele                                                    | tggagcagttctctgtgccaggtctcAcggggctctccggatgcaaggggct          | 60         |
|            |                                                                               |                       |                          | IL28B-Arg32 allele                                                    | tggagcagttctctgtgccaggtctcGggggctctccggatgcaaggggct           | 62         |
| rs629008   | 19                                                                            | 44427130              | intron 2 of IL28B        | mRNA<br>stability/turnover/<br>alternate splicing                     | cttcaggaaacatgagtcagtcctc[A/G]cagtaggagcatgatatagcccacg       | 64         |
| rs628973   | 19                                                                            | 44427106              | intron 2 of IL28B        | mRNA<br>stability/turnover/<br>alternate splicing                     | gggagatgtagagaccctcttck[A/T]maggaaacatgagtcagtcctcga          | 65         |
| 1,         | Chromosome positions are derived from Hapmap project data release 27.         |                       |                          |                                                                       |                                                               |            |
| 2,         | Gene locations were obtained by scanning $\pm 100$ kb from the associated SNP |                       |                          |                                                                       |                                                               |            |
| HR allele, | Allele associated with higher response to therapy.                            |                       |                          |                                                                       |                                                               |            |
| LR allele, | Allele associated with lower response or no response to therapy.              |                       |                          |                                                                       |                                                               |            |

Table 1 continued  
Summary of SNPs

| SNP        | Chromosome                                                                    | Position <sup>1</sup> | Location <sup>2</sup> | SNP effect                                                               | Sequence comprising SNP                                     | SEQ ID NO: |
|------------|-------------------------------------------------------------------------------|-----------------------|-----------------------|--------------------------------------------------------------------------|-------------------------------------------------------------|------------|
| rs8103142  | 19                                                                            | 44426946              | exon 2 of IL28B       | missense: K74R<br>mutation in IL28B<br>HR allele=T or A<br>(IL28B-Lys74) | tccTggggaagagggcgagcgac[C/T]TgcagTccttcagcagaagcgactct      | 66         |
|            |                                                                               |                       |                       | LR allele=C or G<br>(IL28B-Arg74)                                        | agagtcgtcttcTgaaggactgcAgtgcggctccgcctctccccaaggga          | 67         |
|            |                                                                               |                       |                       | mRNA                                                                     | agagtcgtcttcTgaaggactgcAgtgcggctccgcctctccccaaggga          | 69         |
| rs8102358  | 19                                                                            | 44426852              | intron 3 of IL28B     | stability/turnover/<br>alternate splicing<br>mRNA                        | gtgaaggggccactacagagcaggt[A/G]agcagggcTggggagggcagggggTgggg | 71         |
| rs11881222 | 19                                                                            | 44426763              | intron 3 of IL28B     | stability/turnover/<br>alternate splicing                                | agaggggcacagccagTgtgtcaggt[A/G]ggagcagggagggtagcaggtg       | 72         |
| rs61735713 | 19                                                                            | 44426330              | exon 4 of IL28B       | missense: H160Y<br>mutation in IL28B                                     | ccggggggccgcctccaccattggcTgCaccggcTccagggcccccaaaaag        | 73         |
|            |                                                                               |                       |                       | IL28B-His160                                                             | ccggggggccgcctccaccattggcTgCaccggcTccagggcccccaaaaag        | 74         |
|            |                                                                               |                       |                       | IL28B-Tyr160                                                             | ccggggggccgcctccaccattggcTgTaccggcTccagggcccccaaaaag        | 76         |
| rs62120527 | 19                                                                            | 44426192              | exon 5 of IL28B       | missense: E175K<br>mutation in IL28B                                     | gaagaggtTgaaggTgacagaggcct[C/T]Tgagcagccaggggactcctgtgggg   | 78         |
| 1,         | Chromosome positions are derived from Hapmap project data release 27.         |                       |                       |                                                                          |                                                             |            |
| 2,         | Gene locations were obtained by scanning $\pm 100$ kb from the associated SNP |                       |                       |                                                                          |                                                             |            |
| HR allele, | Allele associated with higher response to therapy.                            |                       |                       |                                                                          |                                                             |            |
| LR allele, | Allele associated with lower response or no response to therapy.              |                       |                       |                                                                          |                                                             |            |

Table 1 continued  
Summary of SNPs

| SNP        | Chromosome | Position <sup>1</sup> | Location <sup>2</sup>             | SNP effect         | Sequence comprising SNP                                             | SEQ ID NO: |
|------------|------------|-----------------------|-----------------------------------|--------------------|---------------------------------------------------------------------|------------|
| rs4803217  | 19         | 44426060              | 3'-end of IL28B                   | IL28B-Glu175       | ccctacaggaggtccctggctgctcGaggccctgctacacctcaacctcttc                | 79         |
|            |            |                       |                                   | IL28B-Lys175       | ccctacaggaggtccctggctgctcAaggccctgctacacctcaacctcttc                | 81         |
| rs8105790  | 19         | 44426060              | 3'-end of IL28B                   | mRNA               | tagcgactgggtgacaaataaag[A/C]caagtggtcaattataataaataat               | 83         |
|            |            |                       |                                   | stability/turnover |                                                                     |            |
|            |            |                       |                                   | mRNA               | ttccctctcgacatcatcctcaagtcctg[C/T]ttctgtggttacatctccgctaaatgatgc    | 84         |
|            |            |                       |                                   | HR allele=T        | ttccctctcgacatcatcctcaagtcctgTttctgtggttacatctccgctaaatgatgc        | 85         |
| rs12980275 | 19         | 44424341              | 1.75 kb distal to 3'-end of IL28B | LR allele=C        | ttccctctcgacatcatcctcaagtcctgCttctgtggttacatctccgctaaatgatgc        | 86         |
|            |            |                       |                                   | mRNA               |                                                                     |            |
|            |            |                       |                                   | stability/turnover | gggtgctgagagaagtcacaaatcctagaac[A/G]gacgtgtctaaatattgcccggggtagcggg | 87         |
|            |            |                       |                                   | HR allele=A        | gggtgctgagagaagtcacaaatcctagaacA gacgtgtctaaatattgcccggggtagcggg    | 88         |
| rs7248668  | 19         | 44435661              | 5'-end of IL28B                   | LR allele=G        | gggtgctgagagaagtcacaaatcctagaacG gacgtgtctaaatattgcccggggtagcggg    | 89         |
|            |            |                       |                                   | expression level   | tggaatggtctcagtcgtagcccaagct[A/G]gagcatagtagtggcacaatcggccactcac    | 159        |
|            |            |                       |                                   | HR allele=G        | tggaatggtctcagtcgtagcccaagctG gacatagtagtggcacaatcggccactcac        | 160        |
|            |            |                       |                                   | LR allele=A        | tggaatggtctcagtcgtagcccaagctA gacatagtagtggcacaatcggccactcac        | 161        |

1. Chromosome positions are derived from Hapmap project data release 27.

2. Gene locations were obtained by scanning  $\pm 100$  kb from the associated SNP

HR allele, Allele associated with higher response to therapy.

LR allele, Allele associated with lower response or no response to therapy.

Table 2  
Characteristic for higher responder (HR) and lower-responder (LR) subjects

| Study                                  | Stage One          |                         |                 |                | Stage Two       |                |                 |                |                 |                |                   |               |
|----------------------------------------|--------------------|-------------------------|-----------------|----------------|-----------------|----------------|-----------------|----------------|-----------------|----------------|-------------------|---------------|
|                                        | Australian (n=293) |                         | Berlin (n=298)  |                | Turin (n=93)    |                | UK (n=91)       |                | Bonn (n=43)     |                | Australian (n=32) |               |
|                                        | HR (131)           | LR (162)                | HR (143)        | LR (155)       | HR (50)         | LR (43)        | HR (43)         | LR (48)        | HR (13)         | LR (30)        | HR (13)           | LR (19)       |
| <b>Demographic factors<sup>a</sup></b> |                    |                         |                 |                |                 |                |                 |                |                 |                |                   |               |
| Age (yr)                               | 42.0<br>(10.0)     | 43.9<br>(7.0)           | 41.3<br>(10.4)  | 46.9<br>(10.1) | 43.1<br>(13.0)  | 44.3<br>(10.1) | 40.0<br>(11.4)  | 48.0<br>(11.8) | 39.2<br>(12.8)  | 52.0<br>(10.9) | 34.8<br>(9.9)     | 50.8<br>(4.9) |
| No. Females (%)                        | 51 (38.9)          | 35 (21.6)               | 76 (53.1)       | 67 (43.2)      | 27 (54.0)       | 15 (34.9)      | 14 (32.6)       | 12 (25.0)      | 6 (46.2)        | 10 (33.3)      | 6 (46.2)          | 6 (31.6)      |
| No. Males (%)                          | 80 (61.1)          | 127 (78.4) <sup>c</sup> | 67 (46.9)       | 88 (56.8)      | 23 (46.0)       | 28 (65.1)      | 29 (67.4)       | 36 (75.0)      | 7 (53.8)        | 20 (66.7)      | 7 (53.8)          | 13 (68.4)     |
| BMI                                    | 26.9 (5.0)         | 27.5 (5.1)              | 25.2 (4.5)      | 25.9 (3.9)     | 24.1 (3.4)      | 24.6 (3.5)     | 24.6 (5.9)      | 26.4 (6.4)     | 24.3 (3.5)      | 27.3 (4.6)     | 26.7 (5.3)        | 25.7 (6.3)    |
| Viral load<br>(IU/ml)                  | NS <sup>b</sup>    |                         | NS <sup>b</sup> |                | NS <sup>b</sup> |                | NS <sup>b</sup> |                | NS <sup>b</sup> |                | NS <sup>b</sup>   |               |

a, Unless otherwise specified, mean (s.d.) are presented.

b, No significance within each cohort for chi squared comparison of viral load among R vs NR. This methodology was chosen as the viral titers were measured using different kits with different sensitivity between cohorts.

c, P < 0.05. No significant difference was observed between stages one and two or between cohorts for age, BMI and viral load.

Table 3

Preferred SNPs having alleles associated with efficacy of therapy

| SNP        | Chromosome | Position <sup>1</sup> | Location <sup>2</sup>         | SNP effect                                     | Sequence comprising SNP                                                                                                                                                                                    | SEQ ID NO:        |
|------------|------------|-----------------------|-------------------------------|------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------|
| rs10853728 | 19         | 44436986              | IL28A/IL28B intergenic region | Weak                                           | tgtctcgttaagcagccctggggagatgtgggc[C/G]taagcttgggtgaggatgagagatcgtctt                                                                                                                                       | 3                 |
| rs7248668  | 19         | 44435661              | 5'-end of IL28B               | expression level<br>HR allele=G<br>LR allele=A | tgacatgggtctcagtcgtgtagcccaagct[A/G]gagcatagtagtgccacaaatcgccactcac<br>tgacatgggtctcagtcgtgtagcccaagctGgagcatagtagtgccacaaatcgccactcac<br>tgacatgggtctcagtcgtgtagcccaagctA gagcatagtagtgccacaaatcgccactcac | 159<br>160<br>161 |
| rs8099917  | 19         | 44435005              | 5'-end of IL28B               | expression level<br>HR allele=T<br>LR allele=G | cctcctttgtttcttctcgtgtagcaat[G/T]caccccaaatgggaaccatgctgtatatacag<br>cctcctttgtttcttctcgtgtagcaatTcaccccaaatgggaaccatgctgtatatacag<br>cctcctttgtttcttctcgtgtagcaatGcaccccaaatgggaaccatgctgtatatacag        | 4<br>5<br>6       |
| rs8109886  | 19         | 44434603              | 5'-end of IL28B               | expression level<br>HR allele=C<br>LR allele=A | ttcttatcatttttccacaagaagcatccig[A/C]cccaggctcgtctgtctgtctcaatcaatc<br>ttcttatcatttttccacaagaagcatccigCcccaggctcgtctgtctgtctcaatcaatc<br>ttcttatcatttttccacaagaagcatccigAcccaggctcgtctgtctgtctcaatcaatc     | 9<br>10<br>11     |
| rs4803221  | 19         | 44430969              | 5'-end of IL28B               | expression level<br>HR allele=C<br>LR allele=G | cagcgtccggggctccagcggagcggtagtg[C/G]gagagcagggcagcgcggggggccctctgc<br>cagcgtccggggctccagcggagcggtagtgCgagagcagggcagcgcggggggccctctgc<br>cagcgtccggggctccagcggagcggtagtgGgagagcagggcagcgcggggggccctctgc     | 41<br>162<br>163  |
| rs12979860 | 19         | 44430627              | 5'-end of IL28B               | expression level<br>HR allele=C<br>LR allele=T | tgtactgaaccaggaggagctcccccgaaggcg[C/T]gaaccagggttgtaattgcactccgctcc<br>tgtactgaaccaggaggagctcccccgaaggcgGgaaccagggttgtaattgcactccgctcc<br>tgtactgaaccaggaggagctcccccgaaggcgTgaaccagggttgtaattgcactccgctcc  | 42<br>164<br>165  |

1, Chromosome positions are derived from Hapmap project data release 27.

2, Gene locations were obtained by scanning  $\pm 100$  kb from the associated SNP

HR allele, Allele associated with higher response to therapy.

LR allele, Allele associated with lower response or no response to therapy.

Table 3 continued

Preferred SNPs having alleles associated with efficacy of therapy

|            |    |          |                                      |                                                                          |                                                                   |    |
|------------|----|----------|--------------------------------------|--------------------------------------------------------------------------|-------------------------------------------------------------------|----|
| rs8103142  | 19 | 44426946 | exon 2 of IL28B                      | missense: K74R<br>mutation in IL28B<br>HR allele=T or A<br>(IL28B-Lys74) | tcttgaggagaggaggaggac[C/T]tgcagtccttcagcagaagcgactct              | 66 |
|            |    |          |                                      | LR allele=C or G<br>(IL28B-Arg74)                                        | agagtcgcttctgcgaaggactgcaAgtcgcgcctccgcctcttccccaggga             | 67 |
|            |    |          |                                      | mRNA                                                                     | agagtcgcttctgcgaaggactgcaGgtgcgcgcctccgcctcttccccaggga            | 69 |
|            |    |          | 1.75 kb distal to 3'-end<br>of IL28B | stability/turnover                                                       | ttcccttcgacatcactccaalgctcgt[C/T]tctgtggtacatcttcgciaatgatgc      | 84 |
| rs8105790  | 19 | 44424341 |                                      | HR allele=T                                                              | ttcccttcgacatcactccaalgctcgtTtctgtggtacatcttcgciaatgatgc          | 85 |
|            |    |          |                                      | LR allele=C                                                              | ttcccttcgacatcactccaalgctcgtCtctgtggtacatcttcgciaatgatgc          | 86 |
|            |    |          | 2.47 kb distal to 3'-end<br>of IL28B | mRNA                                                                     | gggtctgagagagagtcnaattctctagaac[A/G]gacgtgtctaaatattgcccgggtagcgg | 87 |
| rs12980275 | 19 | 44423623 |                                      | stability/turnover                                                       | gggtctgagagagagtcnaattctctagaacAgacgtgtctaaatattgcccgggtagcgg     | 88 |
|            |    |          |                                      | HR allele=A                                                              | gggtctgagagagagtcnaattctctagaacAgacgtgtctaaatattgcccgggtagcgg     | 89 |
|            |    |          |                                      | LR allele=G                                                              | gggtctgagagagagtcnaattctctagaacGgacgtgtctaaatattgcccgggtagcgg     |    |

1, Chromosome positions are derived from Hapmap project data release 27.

2, Gene locations were obtained by scanning  $\pm 100$  kb from the associated SNP

HR allele, Allele associated with higher response to therapy.

LR allele, Allele associated with lower response or no response to therapy.

Table 4  
SNP and genotype associations with efficacy of therapy

| SNP            | Stage one<br>p value <sup>1</sup> | Stage 2<br>p value <sup>1</sup> | Merged<br>p value <sup>2</sup> | OR<br>(95% C.I.) <sup>3</sup> | Possible genotypes           | Tested<br>genotype | Genotype P value      | OR (95% C.I.)    |
|----------------|-----------------------------------|---------------------------------|--------------------------------|-------------------------------|------------------------------|--------------------|-----------------------|------------------|
| Chromosome 19: |                                   |                                 |                                |                               |                              |                    |                       |                  |
| rs4803224      | 5.50 x 10 <sup>-02</sup>          | 0.2                             | 0.77                           |                               | C/C C/G G/G                  |                    |                       |                  |
| rs12980602     | 1.08 x 10 <sup>-02</sup>          | 2.66 x 10 <sup>-02</sup>        | 1.02 x 10 <sup>-03</sup>       |                               | C/C C/T T/T                  |                    |                       |                  |
| rs10853728     | 2.67 x 10 <sup>-03</sup>          | 0.97                            | 7.42 x 10 <sup>-02</sup>       |                               | C/C C/G G/G                  |                    |                       |                  |
| rs8099917      | 7.06 x 10 <sup>-08</sup>          | 9.39 x 10 <sup>-04</sup>        | 9.25 x 10 <sup>-09</sup>       | 1.86 (1.49-2.32)              | G/G G/T T/T                  | G/G                | 0.016                 | 2.03(1.09-3.78)  |
|                |                                   |                                 |                                |                               |                              | G/T                | 6.65x10 <sup>-7</sup> | 2.00(1.51-2.64)  |
|                |                                   |                                 |                                |                               |                              | T/T                | 2.64x10 <sup>-9</sup> | 0.44 (0.33-0.58) |
| rs8113007      |                                   |                                 |                                |                               | A/A A/T T/T                  |                    |                       |                  |
| rs8109889      |                                   |                                 |                                |                               | C/C C/T T/T                  |                    |                       |                  |
| rs8109886      | 1.29 x 10 <sup>-04</sup>          | 3.44 x 10 <sup>-02</sup>        | 1.27 x 10 <sup>-04</sup>       |                               | A/A A/C C/C                  |                    |                       |                  |
| rs61599059     |                                   |                                 |                                |                               | */ *CT CT/CT                 |                    |                       |                  |
| rs34567744     |                                   |                                 |                                |                               | */ *CT CT/CT                 |                    |                       |                  |
| rs10642510     |                                   |                                 |                                |                               | */ *CT *TC CT/TC CT/CT TC/TC |                    |                       |                  |
| rs10643535     |                                   |                                 |                                |                               | **/* **CT CT/CT              |                    |                       |                  |
| rs34593676     |                                   |                                 |                                |                               | **/* **TC TC/TC              |                    |                       |                  |
| rs25122122     |                                   |                                 |                                |                               | */ *T T/T                    |                    |                       |                  |

1, Stage one and stage two p-values are based on allelic comparisons obtained from Haploview.

2, Merged p-values are based on cochrane-armitage trend test results.

3, Odds ratio (OR) and 95% confidence interval (95% C.I.) are based on allelic distributions of SNPs for the combined cohort.

Table 4 continued  
SNP and genotype associations with efficacy of therapy

| SNP        | Stage one<br>p value <sup>1</sup> | Stage 2<br>p value <sup>1</sup> | Merged<br>p value <sup>2</sup> | OR<br>(95% C.I.) <sup>3</sup> | Possible genotypes | Tested genotype | Genotype P value | OR (95% C.I.) |
|------------|-----------------------------------|---------------------------------|--------------------------------|-------------------------------|--------------------|-----------------|------------------|---------------|
| rs35407108 |                                   |                                 |                                |                               | */* */A A/A        |                 |                  |               |
| rs59211796 |                                   |                                 |                                |                               | A/A A/G G/G        |                 |                  |               |
| rs62120529 |                                   |                                 |                                |                               | A/A A/G G/G        |                 |                  |               |
| rs62120528 |                                   |                                 |                                |                               | A/A A/C C/C        |                 |                  |               |
| rs12983038 |                                   |                                 |                                |                               | A/A A/G G/G        |                 |                  |               |
| rs10853727 | 0.72                              | 0.22                            | 0.43                           |                               | C/C C/T T/T        |                 |                  |               |
| rs7254424  |                                   |                                 |                                |                               | A/A A/G G/G        |                 |                  |               |
| rs1549928  |                                   |                                 |                                |                               | A/A A/G G/G        |                 |                  |               |
| rs34347451 |                                   |                                 |                                |                               | */* */A A/A        |                 |                  |               |
| rs35814928 |                                   |                                 |                                |                               | */* */A A/A        |                 |                  |               |
| rs4803222  |                                   |                                 |                                |                               | C/C C/G G/G        |                 |                  |               |
| rs11322783 |                                   |                                 |                                |                               | */* */T T/T        |                 |                  |               |
| rs4803221  |                                   |                                 |                                |                               | C/C C/G G/G        |                 |                  |               |
| rs12979860 |                                   |                                 |                                |                               | C/C C/T T/T        |                 |                  |               |
| rs12971396 |                                   |                                 |                                |                               | C/C C/G G/G        |                 |                  |               |
| rs11672932 |                                   |                                 |                                |                               | C/C C/G G/G        |                 |                  |               |
| rs11882871 |                                   |                                 |                                |                               | A/A A/G G/G        |                 |                  |               |

1, Stage one and stage two p-values are based on allelic comparisons obtained from Haploview.

2, Merged p-values are based on cochrane-armitage trend test results.

3, Odds ratio (OR) and 95% confidence interval (95% C.I.) are based on allelic distributions of SNPs for the combined cohort.

Table 4 continued  
SNP and genotype associations with efficacy of therapy

| SNP        | Stage one<br>p value <sup>1</sup> | Stage 2<br>p value <sup>1</sup> | Merged<br>p value <sup>2</sup> | OR<br>(95% C.I.) <sup>3</sup> | Possible genotypes | Tested genotype | Genotype P value | OR (95% C.I.) |
|------------|-----------------------------------|---------------------------------|--------------------------------|-------------------------------|--------------------|-----------------|------------------|---------------|
| rs56215543 |                                   |                                 |                                | A/A                           | A/G G/G            |                 |                  |               |
| rs12979731 |                                   |                                 |                                | C/C                           | C/T T/T            |                 |                  |               |
| rs2020358  |                                   |                                 |                                | G/G                           | G/T T/T            |                 |                  |               |
| rs34853289 |                                   |                                 |                                | C/C                           | C/T T/T            |                 |                  |               |
| rs8107030  |                                   |                                 |                                | A/A                           | A/G G/G            |                 |                  |               |
| rs41537748 |                                   |                                 |                                | C/C                           | C/T T/T            |                 |                  |               |
| rs59702201 |                                   |                                 |                                | */                            | *A/TAT ATAT/ATAT   |                 |                  |               |
| rs2596806  |                                   |                                 |                                | C/C                           | C/G G/G            |                 |                  |               |
| rs2569377  |                                   |                                 |                                | A/A                           | A/G G/G            |                 |                  |               |
| rs4803219  |                                   |                                 |                                | C/C                           | C/T T/T            |                 |                  |               |
| rs28416813 |                                   |                                 |                                | C/C                           | C/G G/G            |                 |                  |               |
| rs630388   |                                   |                                 |                                | A/A                           | A/G G/G            |                 |                  |               |
| rs629976   |                                   |                                 |                                | A/A                           | A/G G/G            |                 |                  |               |
| rs629976   |                                   |                                 |                                |                               |                    | A/A             |                  |               |
| rs629976   |                                   |                                 |                                | A/A                           | A/G G/G            |                 |                  |               |
| rs629008   |                                   |                                 |                                | A/A                           | A/T T/T            |                 |                  |               |
| rs628973   |                                   |                                 |                                |                               |                    | G/G             |                  |               |

1, Stage one and stage two p-values are based on allelic comparisons obtained from Haploview.

2, Merged p-values are based on cochrane-armitage trend test results.

3, Odds ratio (OR) and 95% confidence interval (95% C.I.) are based on allelic distributions of SNPs for the combined cohort.

Table 4 continued  
SNP and genotype associations with efficacy of therapy

| SNP        | Stage one<br>p value <sup>1</sup> | Stage 2<br>p value <sup>1</sup> | Merged<br>p value <sup>2</sup> | OR<br>(95% C.I.) <sup>3</sup> | Possible genotypes | Tested genotype   | Genotype P value   | OR (95% C.I.)                      |
|------------|-----------------------------------|---------------------------------|--------------------------------|-------------------------------|--------------------|-------------------|--------------------|------------------------------------|
| rs8103142  | -                                 | 3.83 x 10 <sup>-04</sup>        | -                              | -                             | C/C C/T T/T        | C/C<br>C/T<br>T/T | 0.0033<br>0.000492 | 0.62(0.39-0.96)<br>2.03(1.36-3.05) |
| rs8102358  | -                                 | -                               | -                              | -                             | A/A A/G G/G        | -                 | -                  | -                                  |
| rs11881222 | -                                 | -                               | -                              | -                             | A/A A/G G/G        | -                 | -                  | -                                  |
| rs61735713 | -                                 | -                               | -                              | -                             | C/C C/T T/T        | -                 | -                  | -                                  |
| rs61735713 | -                                 | -                               | -                              | -                             | -                  | C/C<br>T/T        | -                  | -                                  |
| rs61735713 | -                                 | -                               | -                              | -                             | -                  | -                 | -                  | -                                  |
| rs62120527 | -                                 | -                               | -                              | -                             | C/C C/T T/T        | -                 | -                  | -                                  |
| rs62120527 | -                                 | -                               | -                              | -                             | -                  | C/C<br>T/T        | -                  | -                                  |
| rs62120527 | -                                 | -                               | -                              | -                             | -                  | -                 | -                  | -                                  |
| rs4803217  | -                                 | -                               | -                              | -                             | A/A A/C C/C        | -                 | -                  | -                                  |
| rs8105790  | -                                 | 3.70 x 10 <sup>-04</sup>        | -                              | -                             | C/C C/T T/T        | -                 | -                  | -                                  |

1, Stage one and stage two p-values are based on allelic comparisons obtained from Haploview.

2, Merged p-values are based on cochrane-armitage trend test results.

3, Odds ratio (OR) and 95% confidence interval (95% C.I.) are based on allelic distributions of SNPs for the combined cohort.

Table 4 continued  
SNP and genotype associations with efficacy of therapy

| SNP        | Stage one                |  | Stage 2                  |  | Merged<br>p value <sup>2</sup> | OR<br>(95% C.I.) <sup>3</sup> |  | Possible genotypes | Tested genotype | Genotype P value | OR (95% C.I.)   |
|------------|--------------------------|--|--------------------------|--|--------------------------------|-------------------------------|--|--------------------|-----------------|------------------|-----------------|
|            | p value <sup>1</sup>     |  | p value <sup>1</sup>     |  |                                |                               |  |                    |                 |                  |                 |
| rs12980275 | 4.81 x 10 <sup>-08</sup> |  | 1.24 x 10 <sup>-04</sup> |  | 7.74 x 10 <sup>-10</sup>       |                               |  | A/A A/G G/G        | A/A             | 0.0000908        | 2.06(1.43-2.97) |
|            |                          |  |                          |  |                                |                               |  |                    | A/G             |                  |                 |
|            |                          |  |                          |  |                                |                               |  |                    | G/G             | 0.036            | 0.61(0.39-0.97) |

- 1, Stage one and stage two p-values are based on allelic comparisons obtained from Haploview.
- 2, Merged p-values are based on cochrane-armitage trend test results.
- 3, Odds ratio (OR) and 95% confidence interval (95% C.I.) are based on allelic distributions of SNPs for the combined cohort.

Table 5  
Associations of chromosome 19 SNP combinations with efficacy of therapy

| SNP combination | Possible genotype combinations                                                                                                                                                                                                                                                | Tested genotype combination | Genotype combination | P value   | OR (95% C.I.)   | Sequence comprising SNP (SEQ ID NO:)                                        |
|-----------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------|----------------------|-----------|-----------------|-----------------------------------------------------------------------------|
| rs 12980275     | GG GG, GG TG, AG GG, GG TT,                                                                                                                                                                                                                                                   | AA TT                       |                      |           |                 | gggtgctgagagagtcacattctctagaacAgacgtgtctaaatattgcggggtagcgg (SEQ ID NO: 88) |
| rs 8099917      | AG TG, AG TT, AA TG, AA TT                                                                                                                                                                                                                                                    | (HR HR)                     |                      | 0.0000613 | 2.11(1.46-3.04) | ccctctttgtttcttctgtgagcaatTtcacccaattggaaacatgctgtatcacag(SEQ ID NO: 5)     |
|                 |                                                                                                                                                                                                                                                                               | GG GG                       |                      | 0.042     | 0.47(0.23-0.99) | gggtgctgagagagtcacattctctagaacGgaagtgtctaaatattgcggggtagcgg (SEQ ID NO: 89) |
|                 |                                                                                                                                                                                                                                                                               | (LR LR)                     |                      |           |                 | ccctctttgtttcttctgtgagcaatGtcacccaattggaaacatgctgtatcacag (SEQ ID NO: 6)    |
| rs8103142       |                                                                                                                                                                                                                                                                               | TT TT                       |                      |           |                 | tcttggggaagagcgggagcggcTtgcagtccttcagcagaagcgactct                          |
| rs 8099917      | CC GG, CC TG, CC TT, CT TG, CT TT, TT TT                                                                                                                                                                                                                                      | (HR HR)                     |                      | 0.000492  | 2.03(1.36-3.05) | (reverse complement of SEQ ID NO: 67)                                       |
|                 |                                                                                                                                                                                                                                                                               | CC GG                       |                      | 0.077     |                 | ccctctttgtttcttctgtgagcaatTtcacccaattggaaacatgctgtatcacag (SEQ ID NO: 5)    |
|                 |                                                                                                                                                                                                                                                                               | (LR LR)                     |                      |           |                 | tcttggggaagagcgggagcggcCtgcagtccttcagcagaagcgactct                          |
|                 |                                                                                                                                                                                                                                                                               |                             |                      |           |                 | (reverse complement of SEQ ID NO: 69)                                       |
|                 |                                                                                                                                                                                                                                                                               |                             |                      |           |                 | ccctctttgtttcttctgtgagcaatGtcacccaattggaaacatgctgtatcacag (SEQ ID NO: 6)    |
| rs 12980275     | AA CC GG, AA CC GT, AA CC TT, AA CT GG, AA CT GT, AA CT TT, AA TT GG, AA TT GT, AA TT TT, AG CC GG, AG CC GT, AG CC TT, AG CT GG, AG CT GT, AG CT TT, AG TT GG, AG TT GT, AG TT TT, GG CC GG, GG CC GT, GG CC TT, GG CT GG, GG CT GT, GG CT TT, GG TT GG, GG TT GT, GG TT TT, | AA TT TT                    |                      | 0.00063   | 2.03(1.36-3.05) | gggtgctgagagagtcacattctctagaacAgacgtgtctaaatattgcggggtagcgg (SEQ ID NO: 88) |
| rs8103142       |                                                                                                                                                                                                                                                                               | (HR HR HR)                  |                      |           |                 | tcttggggaagagcgggagcggcTtgcagtccttcagcagaagcgactct                          |
| rs 8099917      |                                                                                                                                                                                                                                                                               |                             |                      | 0.049     | 0.49(0.23-1.01) | (reverse complement of SEQ ID NO: 67)                                       |
|                 |                                                                                                                                                                                                                                                                               | GG CC GG                    |                      |           |                 | ccctctttgtttcttctgtgagcaatTtcacccaattggaaacatgctgtatcacag (SEQ ID NO: 5)    |
|                 |                                                                                                                                                                                                                                                                               | (LR LR LR)                  |                      |           |                 | gggtgctgagagagtcacattctctagaacGgaagtgtctaaatattgcggggtagcgg (SEQ ID NO: 89) |
|                 |                                                                                                                                                                                                                                                                               |                             |                      |           |                 | tcttggggaagagcgggagcggcCtgcagtccttcagcagaagcgactct                          |
|                 |                                                                                                                                                                                                                                                                               |                             |                      |           |                 | (reverse complement of SEQ ID NO: 69)                                       |
|                 |                                                                                                                                                                                                                                                                               |                             |                      |           |                 | ccctctttgtttcttctgtgagcaatGtcacccaattggaaacatgctgtatcacag (SEQ ID NO: 6)    |

HR, genotype homozygous for HR alleles at designated locus associated with higher response to therapy.

LR, genotype homozygous for LR alleles at designated locus associated with lower response to therapy.

Table 6  
Haplotype effects for six chromosome 19 SNP allele combinations on efficacy of therapy

| SNP haplotype<br>for alleles (a)-(f) <sup>1</sup> | Average frequency<br>in cohort (%) | Frequency in<br>HR <sup>2</sup> (%) | Frequency in<br>LR <sup>3</sup> (%) | Haplotype p value      | OR (95% CI) <sup>4</sup> |
|---------------------------------------------------|------------------------------------|-------------------------------------|-------------------------------------|------------------------|--------------------------|
| ATTCT                                             | 45.2                               | 49.4                                | 41.5                                | $1.2 \times 10^{-03}$  | 1.37 (1.13 - 1.67)       |
| GCCTAG                                            | 25.6                               | 18.8                                | 31.5                                | $3.03 \times 10^{-09}$ | 2.0 (1.58 - 2.50)        |
| GTCCAT                                            | 11.2                               | 10.7                                | 11.7                                | 0.52                   | 1.11 (0.81 - 1.50)       |
| ATTTAT                                            | 10.5                               | 13.0                                | 8.3                                 | $1.9 \times 10^{-03}$  | 1.64 (1.20 - 2.25)       |
| ATCTAT                                            | 2.2                                | 2.4                                 | 2.0                                 | 0.51                   | 1.23 (0.64 - 2.36)       |
| GCCTAT                                            | 1.8                                | 1.4                                 | 2.1                                 | 0.27                   | 1.5 (0.71 - 3.18)        |
| GTTTCT                                            | 1.1                                | 1.4                                 | 0.9                                 | 0.42                   | 0.63 (0.25 - 1.56)       |

- Haplotypes are shown in order from left to right for combinations of the following SNP:
  - rs12980275 for which possible alleles are A or G (SEQ ID NO: 87);
  - rs8105790 for which possible alleles are C or T (SEQ ID NO: 84);
  - rs8103142 for which possible alleles are C or T (SEQ ID NO: 66);
  - rs10853727 for which possible alleles are C or T (SEQ ID NO: 31);
  - rs8109886 for which possible alleles are A or C (SEQ ID NO: 9); and
  - rs8099917 for which possible alleles are G or T (SEQ ID NO: 4).
- Odds ratios of each haplotype were calculated as carriage vs non-carriage of the haplotype.
- HR, subjects having a higher response to therapy.
- LR, subjects having a lower response to therapy.

Table 7  
Predictive values of HR alleles and LR alleles of IL28B-linked SNPs

| SNP and genotype                       | Responders or High-responders | Low-responders or non-responders | Sensitivity | Specificity | PPV (LR) | NPV (HR) |
|----------------------------------------|-------------------------------|----------------------------------|-------------|-------------|----------|----------|
| rs4803221 (LR allele= G; HR allele= C) |                               |                                  |             |             |          |          |
| GG                                     | 3 (1.3%)                      | 23 (7.8)                         | 0.078       | 0.987       | 0.885    | 0.456    |
| CG                                     | 68 (29.6%)                    | 157 (53.6)                       |             |             |          |          |
| CC                                     | 159 (69.1%)                   | 113 (38.6)                       | 0.691       | 0.614       | 0.584    | 0.717    |
| Total                                  | 230 (100%)                    | 293 (100%)                       |             |             |          |          |
| rs7248668 (LR allele= A; HR allele= G) |                               |                                  |             |             |          |          |
| AA                                     | 3 (1.3)                       | 23 (7.6)                         | 0.076       | 0.987       | 0.885    | 0.452    |
| AG                                     | 70 (29.9)                     | 163 (53.8)                       |             |             |          |          |
| GG                                     | 161 (68.8)                    | 117 (38.6)                       | 0.688       | 0.614       | 0.579    | 0.718    |
| Total                                  | 234 (100%)                    | 303 (100%)                       |             |             |          |          |

Table 7 continued  
Predictive values of HR alleles and LR alleles of IL28B-linked SNPs

| SNP and genotype                        | Responders or High-responders | Low-responders or non-responders | Sensitivity | Specificity | PPV (LR) | NPV (HR) |
|-----------------------------------------|-------------------------------|----------------------------------|-------------|-------------|----------|----------|
| rs8099917 (LR allele= G; HR allele= T)  |                               |                                  |             |             |          |          |
| GG                                      | 4 (1.8)                       | 18 (6.4)                         | 0.064       | 0.982       | 0.818    | 0.457    |
| GT                                      | 75 (33.3)                     | 146 (52.1)                       |             |             |          |          |
| TT                                      | 146 (64.9)                    | 116 (41.4)                       | 0.649       | 0.586       | 0.587    | 0.675    |
| Total                                   | 225 (100%)                    | 280 (100%)                       |             |             |          |          |
| rs12979860 (LR allele= T; HR allele= C) |                               |                                  |             |             |          |          |
| TT                                      | 20 (8.7)                      | 58 (25.8)                        | 0.196       | 0.913       | 0.744    | 0.467    |
| TC                                      | 95 (41.5)                     | 177 (78.7)                       |             |             |          |          |
| CC                                      | 114 (49.8)                    | 61 (27.1)                        | 0.498       | 0.794       | 0.651    | 0.671    |
| Total                                   | 229 (100%)                    | 296 (100%)                       |             |             |          |          |

Table 8  
Effects of HLA-C alleles on efficacy of therapy

| HLA-C allele | Responders and high-responders (n=130) | Responders and high-responders (%) | Low-responders and non-responders (n=171) | Low-responders and non-responders (%) |
|--------------|----------------------------------------|------------------------------------|-------------------------------------------|---------------------------------------|
| Cw1          | 8                                      | 3.1                                | 5                                         | 1.5                                   |
| Cw2          | 8                                      | 3.1                                | 15                                        | 4.4                                   |
| Cw3          | 40                                     | 15.4                               | 30                                        | 8.8                                   |
| Cw4          | 23                                     | 8.8                                | 40                                        | 11.7                                  |
| Cw5          | 22                                     | 8.5                                | 38                                        | 11.1                                  |
| Cw6          | 24                                     | 9.2                                | 36                                        | 10.5                                  |
| Cw7          | 82                                     | 31.5                               | 122                                       | 35.7                                  |
| Cw8          | 11                                     | 4.2                                | 15                                        | 4.4                                   |
| Cw12         | 12                                     | 4.6                                | 15                                        | 4.4                                   |
| Cw14         | 0                                      | 0                                  | 3                                         | 0.9                                   |
| Cw15         | 9                                      | 3.5                                | 10                                        | 2.9                                   |
| Cw16         | 21                                     | 8.1                                | 10                                        | 2.9                                   |
| Cw17         | 0                                      | 0                                  | 3                                         | 0.9                                   |

Table 9

Effect of HLA-C genotype on efficacy of therapy

| HLA-C genotype | Responders and high-responders (n=130) | Low-responders and non-responders (n=171) |
|----------------|----------------------------------------|-------------------------------------------|
| C1+/C1+        | 58                                     | 67                                        |
| C1+/C2+        | 58                                     | 66                                        |
| C2+/C2+        | 14                                     | 38                                        |
| Total C1+      | 174                                    | 200                                       |
| Total C2+      | 86                                     | 142                                       |

Table 10

Significance, specificity and positive predictive value (PPV) of HLA-C2 homozygosity for efficacy of therapy

| Genotype | Responders and high-responders (%) | Low-responders and non-responders (%) | P value*                   | Specificity | Positive Predictive Value (PPV) |
|----------|------------------------------------|---------------------------------------|----------------------------|-------------|---------------------------------|
| C1+/C1+  | 58 (45)                            | 67 (40)                               | 0.343                      |             |                                 |
| C1+/C2+  | 58 (45)                            | 66 (39)                               |                            |             |                                 |
| C2+/C2+  | 14 (11)                            | 38 (23)                               | 0.009, OR=2.37 (1.22-4.59) | 0.918       | 0.731                           |
| C1+      | 174 (66.9)                         | 200 (58.5)                            |                            |             |                                 |
| C2+      | 86 (33.1)                          | 142 (41.5)                            | 0.034, OR=1.44 (1.03-2.01) | 0.669       | 0.623                           |

Table 11  
Combined effects of HLA-C genotype and rs8099917 genotype on efficacy of therapy

| Genotype | Responders and high-responders (n) | Low-responders and non-responders (n) | P value   |
|----------|------------------------------------|---------------------------------------|-----------|
| C1-C1 TT | 41                                 | 30                                    |           |
| C1-C2 TT | 35                                 | 20                                    |           |
| C2-C2 TT | 11                                 | 13                                    |           |
|          |                                    |                                       |           |
| C1-C1 GT | 16                                 | 33                                    |           |
| C1-C2 GT | 21                                 | 40                                    |           |
| C2-C2 GT | 3                                  | 22                                    |           |
|          |                                    |                                       |           |
| C1-C1 GG | 1                                  | 4                                     |           |
| C1-C2 GG | 2                                  | 5                                     |           |
| C2-C2 GG | 0                                  | 3                                     | 0.0000486 |

Table 12  
Reproducibility in the effect of HLA-C genotype and the combined effects of HLA-C genotype and rs8099917 genotypes on efficacy of therapy

| Cohort            | Australian (n=301)                       |                                             |                                     | Replication (n=483)                         |                          |                                     | Combined (n=784)                         |                                             |                                     |
|-------------------|------------------------------------------|---------------------------------------------|-------------------------------------|---------------------------------------------|--------------------------|-------------------------------------|------------------------------------------|---------------------------------------------|-------------------------------------|
|                   | Responders and high-responders n=130 (%) | Low-responders and non-responders n=171 (%) | P value <sup>a</sup><br>OR (95% CI) | Low-responders and non-responders n=228 (%) | Non-responders n=255 (%) | P value <sup>a</sup><br>OR (95% CI) | Responders and high-responders n=358 (%) | Low-responders and non-responders n=426 (%) | P value <sup>a</sup><br>OR (95% CI) |
| <b>C2-C2</b>      | 14 (11)                                  | 38 (23)                                     | 6.4 x 10 <sup>-3</sup>              | 33 (14.4)                                   | 44 (17.3)                | 0.23                                | 47 (13.1)                                | 82 (19.3)                                   | 1.4 x 10 <sup>-2</sup>              |
| <b>Others</b>     | 116 (89)                                 | 133 (77)                                    | 2.37 (1.22 - 4.59)                  | 194 (85.6)                                  | 211 (82.7)               |                                     | 310 (86.9)                               | 344 (80.7)                                  | 1.57 (1.06 - 2.32)                  |
| <b>G*</b>         | 43 (33.1)                                | 107 (62.9)                                  | 2.24 x 10 <sup>-7</sup>             | 92 (40.4)                                   | 134 (52.5)               | 4.74 x 10 <sup>-3</sup>             | 135 (37.7)                               | 241 (56.7)                                  | 7.77 x 10 <sup>-8</sup>             |
| <b>TT</b>         | 87 (66.9)                                | 63 (37.1)                                   | 3.43 (2.13 - 5.55)                  | 136 (59.6)                                  | 121 (47.5)               | 1.64 (1.14 - 2.35)                  | 223 (62.3)                               | 184 (43.3)                                  | 2.16 (1.62 - 2.88)                  |
| <b>C2-C2 + G*</b> | 3 (2.3)                                  | 25 (14.7)                                   | 1.16 x 10 <sup>-4</sup>             | 8 (3.5)                                     | 23 (9.0)                 | 1.02 x 10 <sup>-2</sup>             | 11 (3.1)                                 | 48 (11.3)                                   | 6.23 x 10 <sup>-6</sup>             |
| <b>Others</b>     | 127 (97.7)                               | 145 (85.3)                                  | 7.30 (2.15 - 24.7)                  | 220 (96.5)                                  | 232 (91.0)               | 2.73 (1.19 - 6.22)                  | 347 (96.9)                               | 377 (88.7)                                  | 4.02 (2.05 - 7.86)                  |

Table 13  
Predictive values of HR alleles and LR alleles of IL28B-linked SNPs when combined with HLA-C genotype

| Marker combination and genotype | Responders | Low-responders and non-responders | Representation in sample (%) | Sensitivity | Specificity | PPV (LR + HLA-C2) | NPV (HR+HLA-C1) |
|---------------------------------|------------|-----------------------------------|------------------------------|-------------|-------------|-------------------|-----------------|
| HLA-C and rs4803221             |            |                                   |                              |             |             |                   |                 |
| C2-C2 GG                        | 0          | 3 (1.2)                           | 0.6                          | 0.011       | 1.0         | 1.0               | 0.450           |
| C2-C2 GC                        | 7 (3.4)    | 33 (12.8)                         | 8.6                          | 0.13        | 0.967       | 0.825             | 0.477           |
| C2-C2 G*                        | 7 (3.4)    | 36 (14)                           | 9.3                          | 0.139       | 0.966       | 0.837             | 0.475           |
| C1* G*                          | 58 (27.8)  | 129 (50)                          | 40.1                         | 0.392       | 0.721       | 0.690             | 0.428           |
| Others                          | 143 (68.8) | 93 (36.0)                         | 50.6                         | 0.687       | 0.639       | 0.606             | 0.717           |
| Total                           | 208 (100%) | 258 (100%)                        | 100%                         |             |             |                   |                 |
| HLA-C and rs7248668             |            |                                   |                              |             |             |                   |                 |
| C2-C2 AA                        | 0          | 3 (1.1)                           | 0.6                          | 0.011       | 1.0         | 1.0               | 0.446           |
| C2-C2 AG                        | 7 (3.3)    | 36 (13.6)                         | 9.0                          | 0.135       | 0.967       | 0.837             | 0.471           |
| C2-C2 A*                        | 7 (3.4)    | 39 (14.8)                         | 9.6                          | 0.147       | 0.967       | 0.848             | 0.474           |
| C1* A*                          | 60 (28.4)  | 129 (48.6)                        | 39.8                         | 0.487       | 0.716       | 0.620             | 0.526           |
| Others                          | 144 (68.2) | 97 (36.6)                         | 50.6                         | 0.682       | 0.634       | 0.597             | 0.715           |
| Total                           | 211 (100%) | 265 (100%)                        | 100%                         |             |             |                   |                 |

Table 13 continued  
 Predictive values of HR alleles and LR alleles of IL28B-linked SNPs when combined with HLA-C genotype

| Marker combination and genotype | Responders | Low-responders and non-responders | Representation in sample (%) | Sensitivity | Specificity | PPV (LR + HLA-C2) | NPV (HR+HLA-C1) |
|---------------------------------|------------|-----------------------------------|------------------------------|-------------|-------------|-------------------|-----------------|
| HLA-C and rs8099917             |            |                                   |                              |             |             |                   |                 |
| C2-C2 GG                        | 1 (0.5)    | 3 (1.2)                           | 0.9                          | 0.011       | 0.495       | 0.750             | 0.450           |
| C2-C2 GT                        | 7 (3.3)    | 34 (13.2)                         | 8.8                          | 0.132       | 0.966       | 0.829             | 0.475           |
| C2-C2 G*                        | 8 (3.8)    | 37 (14.4)                         | 9.7                          | 0.144       | 0.962       | <b>0.822</b>      | 0.477           |
| C1* G*                          | 69 (33)    | 113 (44)                          | 39                           | 0.440       | 0.670       | 0.621             | 0.493           |
| Others                          | 132 (63.2) | 107 (41.6)                        | 51.3                         | 0.631       | 0.584       | 0.552             | 0.661           |
| Total                           | 209 (100%) | 257 (100%)                        | 100%                         |             |             |                   |                 |
| HLA-C/ rs12979860               |            |                                   |                              |             |             |                   |                 |
| C2-C2 TT                        | 2 (0.9)    | 8 (3.0)                           | 2.1                          | 0.030       | 0.990       | <b>0.800</b>      | 0.450           |
| C2-C2 TC                        | 12 (5.7)   | 41 (15.6)                         | 11.2                         | 0.156       | 0.943       | 0.774             | 0.473           |
| C2-C2 T*                        | 14 (6.6)   | 49 (18.6)                         | 13.3                         | 0.186       | 0.934       | <b>0.778</b>      | 0.479           |
| C1* T*                          | 94 (44.6)  | 163 (62)                          | 54.2                         | 0.620       | 0.555       | 0.634             | 0.539           |
| Others                          | 103 (48.8) | 51 (19.4)                         | 32.5                         | 0.488       | 0.806       | 0.669             | 0.663           |
| Total                           | 211 (100%) | 263 (100%)                        | 100%                         |             |             |                   |                 |

Table 14  
Demographic characteristics of chronic hepatitis C patients after therapy

|                         | Australian Cohort<br>(n=312313) |                               | Berlin Cohort<br>(n=310307) |                    | Newcastle, UK Cohort<br>(n=6990) |                              | Bonn Cohort<br>(n=57) |                | Trent UK Cohort<br>(n=4843) |                      | Turin Cohort<br>(n=11495) |                     | Total Cohort<br>(n=910905)         |                                              |
|-------------------------|---------------------------------|-------------------------------|-----------------------------|--------------------|----------------------------------|------------------------------|-----------------------|----------------|-----------------------------|----------------------|---------------------------|---------------------|------------------------------------|----------------------------------------------|
|                         | SVR                             | NSVR                          | SVR                         | NSVR               | SVR                              | NSVR                         | SVR                   | NSVR           | SVR                         | NSVR                 | SVR                       | NSVR                | SVR                                | NSVR                                         |
|                         | (n=130)                         | (n=182183)                    | (n=150149)                  | (n=15860)          | (n=3140)                         | (n=5038)                     | (n=26)                | (n=31)         | (n=2221)                    | (n=2622)             | (n=5845)                  | (n=5650)            | (n=417411)                         | (n=403494)                                   |
| Age (yr) -<br>mean (SD) | 40.0 <sup>a</sup><br>(9.6)      | 44.54 <sup>a</sup><br>(7.12)  | 41.0<br>(10.5)              | 46.78<br>(10.34)   | 38.72 <sup>a</sup><br>(11.86)    | 46.0 <sup>a</sup><br>(12.10) | 44.7<br>(12.9)        | 50.8<br>(10.9) | 39.843.5<br>(9.80)          | 45.747.4<br>(7.99.3) | 41.63.3<br>(13.10)        | 45.17<br>(10.09.7)  | 40.96 <sup>a</sup><br>(10.87)      | 45.7 <sup>a</sup> 8 <sup>a</sup><br>(9.32)   |
| Gender - N              |                                 |                               |                             |                    |                                  |                              |                       |                |                             |                      |                           |                     |                                    |                                              |
| Females                 | 52 <sup>b</sup><br>(40.0)       | 42.43 <sup>b</sup><br>(23.15) | 79.78<br>(52.73)            | 69.68<br>(43.10)   | 9.12<br>(29.030.0)               | 10.12<br>(26.324.0)          | 11<br>(42.3)          | 11<br>(35.5)   | 6.5<br>(27.323.8)           | 5.4<br>(1918.2)      | 28.18<br>(48.340.0)       | 19<br>(33.938.0)    | 185.176 <sup>b</sup><br>(44.442.8) | 156 <sup>b</sup> 157 <sup>b</sup><br>(31.68) |
| Males                   | 78 <sup>b</sup><br>(60.0)       | 140 <sup>b</sup><br>(76.95)   | 91.90<br>(47.37)            | 91.90<br>(55.97.0) | 22.28<br>(7170.0)                | 28.38<br>(73.776.0)          | 15<br>(57.7)          | 20<br>(64.5)   | 16<br>(72.776.2)            | 21.18<br>(8081.8)    | 30.27<br>(51.760.0)       | 37.31<br>(66.162.0) | 232.235 <sup>b</sup><br>(55.657.2) | 337 <sup>b</sup><br>(68.42)                  |
| BMI - mean<br>(SD)      | 26.927.0<br>(4.85.1)            | 27.4<br>(5.32)                | 25.1<br>(4.5)               | 25.9<br>(3.9)      | 25.33.7<br>(3.96.3)              | 26.227.0<br>(6.64.8)         | 25.4<br>(4.23)        | 27.3<br>(4.6)  | 26.97.1<br>(3.57)           | 25.26.0<br>(2.93.5)  | 24.023.9<br>(3.23)        | 24.56<br>(3.34)     | 25.57<br>(4.75)                    | 26.35<br>(4.76)                              |
| Viral load <sup>c</sup> | p=NS0.24                        |                               | p < 0.05001                 |                    | p < =0.05002                     |                              | p <                   |                | NSp=0.58                    |                      | p < =0.0502               |                     | p < 0.05001                        |                                              |

<sup>a</sup> p < 0.001 comparisons between responders (SVR) and non-responders (NSVR) based on Student's t-test Unless otherwise specified, mean (s.d.) are presented.

<sup>b</sup> p < 0.05 005 comparisons between responders (SVR) and non-responders (NSVR) based on the  $\chi^2$  test.

<sup>c</sup> Comparisons between SVR and NSVR based on the Mann-Whitney U test.

Table 15  
Table of Haplotypes with ORs

| No. | Haplotype     | Haplotype<br>tagging<br>SNP <sup>a</sup> | Frequency<br>(%) | Responders<br>(%) | Non-<br>responders<br>(%) | P value                | OR <sup>b</sup><br>95% CI |
|-----|---------------|------------------------------------------|------------------|-------------------|---------------------------|------------------------|---------------------------|
| 1   | ATTGATCCTCTG  | rs8109886                                | 43.2             | 48.4              | 39.6                      | $4.00 \times 10^{-4}$  | 0.70,<br>0.57 - 0.85      |
| 2   | GCCAGCTGTAGA  | rs8099917/<br>rs7248668                  | 23.8             | 16.5              | 30.4                      | $9.51 \times 10^{-11}$ | 2.20,<br>1.72 - 2.80      |
| 3   | GCTAGCTCCATG  | rs10853727                               | 10.3             | 10.1              | 10.6                      | 0.77                   | 1.04,<br>0.75 - 1.43      |
| 4   | ATTGATCCTATG  |                                          | 9.8              | 12.3              | 7.8                       | $2.40 \times 10^{-3}$  | 0.60,<br>0.43 - 0.83      |
| 5   | ATTAACCTCTATG |                                          | 1.9              | 2.2               | 1.7                       | 0.47                   | 0.79,<br>0.39 - 1.60      |
| 6   | GCCAGCTGTATG  |                                          | 1.4              | 1.1               | 1.6                       | 0.39                   | 1.49,<br>0.62 - 3.56      |

The distribution of the six haplotypes in haplotype block 2 (Fig 4) bound by SNPs

rs12980275 and rs7248668. The order is: rs12980275, rs12982533, rs8105790,

rs688187, rs11881222, rs8103142, rs12979860, rs48031221, rs10853727, rs8109886,

rs8099917, rs7248668.

<sup>a</sup>SNP in the haplotype block 2 that tag the respective haplotypes:

<sup>b</sup>ORs have been calculated as carriage of the haplotype vs non-carriage of the haplotype.

Table 16

Prediction of failure to clear of virus on therapy with PegIFN/R with homozygote non-responder SNPs  
(based on 404 responders and 464 non responders of European origin)

| Genotype      | Sensitivity<br>(95% CI) | Specificity<br>(95% CI) | Positive                     |                              | Negative                     |                              |
|---------------|-------------------------|-------------------------|------------------------------|------------------------------|------------------------------|------------------------------|
|               |                         |                         | Predictive<br>Value (95% CI) | Predictive<br>Value (95% CI) | Predictive<br>Value (95% CI) | Predictive<br>Value (95% CI) |
| rs4803221 GG  | 8.4<br>(6.0, 11.4)      | 97.0<br>(94.7, 98.5)    | 77.1<br>(62.7, 88.0)         | 47.1<br>(43.5, 50.7)         |                              |                              |
| rs7248668 AA  | 8.1<br>(5.7, 11.0)      | 97.3<br>(95.1, 98.7)    | 78.3<br>(63.6, 89.1)         | 47.0<br>(43.4, 50.5)         |                              |                              |
| rs8099917 GG  | 7.4<br>(5.2, 10.3)      | 96.8<br>(94.5, 98.3)    | 73.3<br>(58.1, 85.4)         | 46.8<br>(43.2, 50.4)         |                              |                              |
| rs12979860 TT | 18.9<br>(15.4, 23.0)    | 89.5<br>(85.9, 92.5)    | 68.3<br>(59.2, 76.5)         | 48.0<br>(44.2, 51.8)         |                              |                              |

## WE CLAIM:

1. A method for accurately determining the likelihood that a subject will or will not respond to treatment with an immunomodulatory composition, said method comprising  
5 detecting homozygosity of a HLA-C2 allele and at least one IL28B low response (LR) allele in a sample from the subject, wherein the IL28B LR allele is the LR allele of rs7248668, and wherein detection of said homozygosity and said at least one LR allele is indicative of a low response of the subject to treatment with said composition.
- 10 2. A method for accurately determining the likelihood that a subject will or will not respond to treatment with an immunomodulatory composition, said method comprising detecting homozygosity of HLA-C2 and rs7248668-LR alleles in a sample from the subject, wherein detection of the double homozygosity is indicative of a low response of the subject to treatment with said composition.
- 15 3. A method for accurately determining the likelihood that a subject will or will not respond to treatment with an immunomodulatory composition, said method comprising detecting two or more markers in a sample from the subject, wherein at least one marker is linked to an HLA-C2 allele and at least one marker is linked to the low  
20 response (LR) allele of rs7248668, and wherein said detection is indicative of HLA-C2 homozygosity and at least one LR allele of rs7248668 in the genome of the subject, said indication being further indicative of a low response of the subject to treatment with said composition.
- 25 4. The method according to any one of claims 1 to 3, wherein the rs7248668-LR allele comprises adenine or adenosine at a position corresponding to position 31 of SEQ ID NO: 161 in the context of said sequence or a shorter sequence comprising at least about 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or  
30 19 or 20 contiguous nucleotides of SEQ ID NO: 161 including the adenine or adenosine at position 31 thereof, a sequence complementary thereto.

5. The method according to any one of claims 1 to 3, wherein the HLA-C2 allele is selected from HLA-Cw2, HLA-Cw4, HLA-Cw5, HLA-Cw6, HLA-Cw15, HLA-Cw17, HLA-Cw18, HLA-Cw\*0114, HLA-Cw\*0307, HLA-Cw\*0315, HLA-Cw\*0345, HLA-Cw\*0707, HLA-Cw\*0709, HLA-Cw\*0776, HLA-Cw\*0810, HLA-Cw\*1204, HLA-Cw\*1205, HLA-Cw\*1209, HLA-Cw\*1221, HLA-Cw\*1233, HLA-Cw\*1241, HLA-Cw\*1412, HLA-Cw\*1602, HLA-Cw\*1609, HLA-Cw\*1612, HLA-Cw\*1619, and HLA-Cw\*1625, and being other than HLA-Cw\*0227 and HLA-Cw\*0411, HLA-Cw\*0429, HLA-Cw\*0436, HLA-Cw\*0455, HLA-Cw\*0611, HLA-Cw\*1507, HLA-Cw\*1525, or HLA-Cw\*1543.
6. The method according to any one of claims 1 to 3 further comprising detecting one or more additional low response (LR) alleles of IL28B in a sample from the subject, wherein detection of HLA-C2 homozygosity and at least one rs7248668-LR allele and at least one additional LR allele of IL28B is indicative of a low response of the subject to treatment with said composition.
7. The method according to claim 6, wherein the at least one additional low response (LR) allele of IL28B is selected from the group of single nuclear polymorphisms (SNPs) set forth in Table 1 or comprises a SNP set forth in Table 1 or is encoded by nucleic acid comprising a SNP set forth in Table 1 or linked to a SNP set forth in Table 1.
8. The method according to claim 7, wherein the at least one additional low response (LR) allele of IL28B is the rs4803221-LR allele.
9. The method according to claim 7, wherein the at least one additional low response (LR) allele of IL28B is the rs8099917-LR allele.
10. The method according to claim 7, wherein the at least one additional low response (LR) allele of IL28B is the rs12979860-LR allele.

11. The method according to any one of claims 6 to 10 comprising detecting two or more additional IL28B-LR alleles at different loci of IL28B.
12. The method according to claim 11, wherein the two or more additional IL28B-LR  
5 alleles are selected from rs8099917-LR, rs4803221-LR and rs12979860-LR.
13. The method according to any one of claims 6 to 10 comprising detecting a haplotype comprising an additional IL28B-LR allele.
- 10 14. The method according to claim 13, wherein the haplotype comprises at least one LR allele of rs4803221, rs12979860, or rs8099917.
- 15 15. The method according to claim 13, wherein the haplotype comprises alleles at rs12980275, rs8105790, rs8103142, rs10853727, rs8109886 and rs8099917.
16. The method according to any one of claims 6 to 15 comprising detecting heterozygosity of an additional IL28B-LR allele(s).
17. The method according to any one of claims 6 to 15 comprising detecting  
20 homozygosity of an additional IL28B-LR allele(s).
18. The method according to any one of claims 6 to 15 comprising detecting at least two additional IL28B-LR alleles at different loci of the IL28B gene.
- 25 19. The method according to any one of claims 6 to 15 comprising detecting at least three additional IL28B-LR alleles at different loci of the IL28B gene.
20. The method according to any one of claims 6 to 15 comprising detecting at least four additional IL28B-LR alleles at different loci of the IL28B gene.

21. The method according to any one of claims 6 to 15 comprising detecting at least five additional IL28B-LR alleles at different loci of the IL28B gene.
22. The method according to any one of claims 6 to 15 comprising detecting at least  
5 six additional IL28B-LR alleles at different loci of the IL28B gene.
23. The method according to any one of claims 6 to 15 comprising detecting at least seven additional IL28B-LR alleles at different loci of the IL28B gene.
- 10 24. The method according to any one of claims 6 to 15 comprising detecting at least eight additional IL28B-LR alleles at different loci of the IL28B gene.
25. The method according to any one of claims 17 to 23, wherein the additional IL28B-LR alleles at different loci of the IL28B gene are selected from the LR allele of  
15 rs4803221, the LR allele of rs12979860, the LR allele of rs8099917, the LR allele of rs12980275, the LR allele of rs8105790, the LR allele of rs8103142, the LR allele of rs10853727, the LR allele of rs8109886 and the LR allele of rs8099917.
26. The method according to any one of claims 1 to 24 comprising performing a  
20 nucleic acid-based assay to thereby detect the HLA-C2 and/or rs7248668-LR allele and/or other optional IL28B-LR allele(s).
27. The method according to any one of claims 1 to 24 comprising performing a nucleic acid-based assay to thereby detect the HLA-C2 and rs7248668-LR allele and  
25 optionally, other IL28B-LR allele(s).
28. The method according to claim 25 or 26 comprising performing an amplification reaction to thereby synthesize nucleic acid specific for the HLA-C2 and/or rs7248668-LR allele and/or other optional IL28B-LR allele(s) of the subject.

29. The method according to any one of claims 1 to 24 comprising performing an antigen-based assay to thereby detect the HLA-C2 and/or rs7248668-LR allele and/or other optional IL28B-LR allele(s).
- 5 30. The method according to any one of claims 1 to 24 comprising performing an antigen-based assay to thereby detect the HLA-C2 and rs7248668-LR allele and optionally, other IL28B-LR allele(s).
31. The method according to claim 28 or 29 comprising contacting a biological sample  
10 derived from a subject with an antibody or ligand capable of specifically binding to an allelic variant of a HLA-C protein and/or an antibody or ligand capable of specifically binding to an allelic variant of rs7248668 and/or optionally, one or more other antibodies or ligands each being capable of specifically binding to a HR or LR allelic variant of IL28B protein for a time and under conditions sufficient for complex to form  
15 and then detecting the complex.
32. The method according to any one of claims 1 to 27, wherein the sample comprises a nucleated cell and/or an extract thereof.
- 20 33. The method according to any one of claims 1 to 30, wherein the sample is selected from the group consisting of whole blood, serum, plasma, peripheral blood mononuclear cells (PBMC), a buffy coat fraction, saliva, urine, a buccal cell, liver biopsy and a skin cell.
- 25 34. The method according to any one of claims 1 to 32, wherein the sample has been derived or isolated or obtained previously from the subject.
35. The method according to any one of claims 1 to 27 or 31, wherein the sample comprises genomic DNA, mRNA, protein or a derivative thereof.

36. The method according to claim 34, wherein the derivative comprises amplified DNA or cDNA.
37. The method according to any one of claims 1 to 35, wherein the subject is  
5 Caucasian.
38. The method according to any one of claims 1 to 35, wherein the subject is African or Asian.
- 10 39. The method according to any one of claims 1 to 37, wherein the immunomodulatory composition comprises one or more IFNs and/or one or more derivatives of said one or more of said IFNs.
40. The method according to claim 38, wherein the composition comprises one or  
15 more IFNs selected from IFN- $\alpha$ , IFN- $\beta$ , IFN- $\omega$ , IFN- $\gamma$ , IFN- $\lambda$ 1, IFN- $\lambda$ 2 and IFN- $\lambda$ 3 and/or one or more derivatives of any one or more of said IFNs.
41. The method according to any one of claims 1 to 37, wherein the immunomodulatory composition comprises one or more guanosine analogs and/or one  
20 or more derivatives of said one or more of said guanosine analogs.
42. The method according to claim 40, wherein the composition comprises one or more guanosine analogs selected from ribavirin, viramidine, 7-benzyl-8-bromoguanine, 9-benzyl-8-bromoguanine, and CpG-containing oligonucleotide(s), and derivative(s),  
25 salt(s), solvate(s) and hydrate(s) thereof.
43. The method according to any one of claims 1 to 37, wherein the immunomodulatory composition comprises IFN- $\alpha$  and ribavirin.
- 30 44. The method according to claim 42, wherein the IFN is pegylated IFN.

45. A process comprising:
- (i) performing a method according to any one of claims 1 to 43; and
  - (ii) administering or recommending an immunomodulatory composition to a subject.
- 5 46. A process comprising:
- (i) obtaining results of a method according to any one of claims 1 to 43; and
  - (ii) administering or recommending an immunomodulatory composition to a subject.
47. A process for selecting a subject in need of treatment with an immunomodulatory  
10 composition, said process comprising:
- (i) exposing a sample comprising cells obtained from the subject to the immunomodulatory composition *in vitro*; and
  - (ii) performing the method according to any one of claims 1 to 43 to thereby identify a subject likely to respond to treatment with the immunomodulatory composition; and
  - 15 (iii) administering or recommending an immunomodulatory composition to a subject likely to respond to treatment.
48. A process for selecting a subject in need of treatment with an immunomodulatory composition, said process comprising:
- 20 (i) exposing a sample comprising cells obtained from the subject to the immunomodulatory composition *in vitro*; and
  - (ii) performing the method according to any one of claims 1 to 43 to thereby identify a subject likely to not respond to treatment with the immunomodulatory composition or likely to provide a low response to treatment; and
  - 25 (iii) administering or recommending an alternative therapy to the immunomodulatory composition.
49. The process of claim 46 or 47, wherein the subject is infected with HCV.
- 30 50. The process of claim 46 or 47, wherein the cells are peripheral blood mononuclear cells or other source of DNA.

51. The process according to any one of claims 44 to 49, wherein the immunomodulatory composition comprises one or more IFNs and/or one or more derivatives of said one or more of said IFNs.
- 5
52. The process according to claim 49, wherein the composition comprises one or more IFNs selected from IFN- $\alpha$ , IFN- $\beta$ , IFN- $\omega$ , IFN- $\gamma$ , IFN- $\lambda$ 1, IFN- $\lambda$ 2 and IFN- $\lambda$ 3 and/or one or more derivatives of any one or more of said IFNs.
- 10 53. The process according to any one of claims 44 to 49, wherein the immunomodulatory composition comprises one or more guanosine analogs and/or one or more derivatives of said one or more of said guanosine analogs.
54. The process according to claim 52, wherein the composition comprises one or  
15 more guanosine analogs selected from ribavirin, viramidine, 7-benzyl-8-bromoguanine, 9-benzyl-8-bromoguanine, and CpG-containing oligonucleotide(s), and derivative(s), salt(s), solvate(s) and hydrate(s) thereof.
55. The process according to any one of claims 44 to 49, wherein the  
20 immunomodulatory composition comprises IFN- $\alpha$  and ribavirin.
56. The process according to any one of claims 50, 51 or 54, wherein the IFN is pegylated IFN.
- 25 57. The process according to any one of claims 46 to 49, wherein the cells are obtained from the subject prior to any *in vivo* administration of the immunomodulatory composition to the subject.
58. The process according to any one of claims 44 to 49, wherein the subject in need is  
30 a subject who has received prior *in vivo* administration of the immunomodulatory composition and wherein the treatment is continued treatment.

59. A process for treating an HCV-infected subject, comprising performing the process according to any one of claims 1 to 43 on a sample from said subject and administering or recommending a therapeutically effective amount of an immunomodulatory composition comprising an IFN to the subject if the subject is likely to respond to treatment or administering or recommending an alternative therapy if the subject is not likely to respond to treatment or likely to produce a low response to treatment.
- 10 60. A process for determining a predisposition in a subject to a chronic HCV infection, said process comprising performing the method according to any one of claims 1 to 43 to thereby identify a subject likely to not respond to treatment with an immunomodulatory composition or likely to provide a low response to treatment, and determining that the subject has a predisposition to chronic HCV infection.
- 15 61. A kit comprising isolated nucleic acid or antibodies for performing a method according to any one of claims 1 to 43.
62. Use of isolated nucleic acid or antibodies in the manufacture of a kit or solid substrate for performing a method according to any one of claims 1 to 43.
- 20

1/9

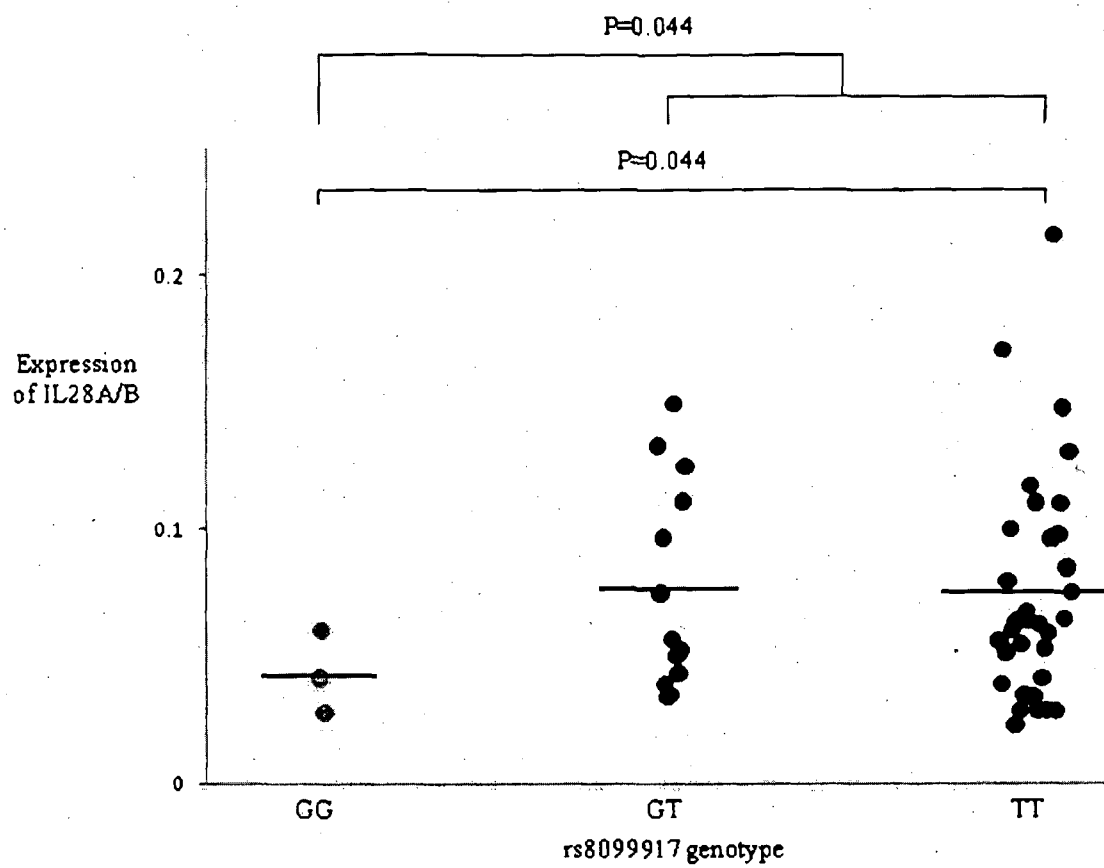


Figure 1

2/9

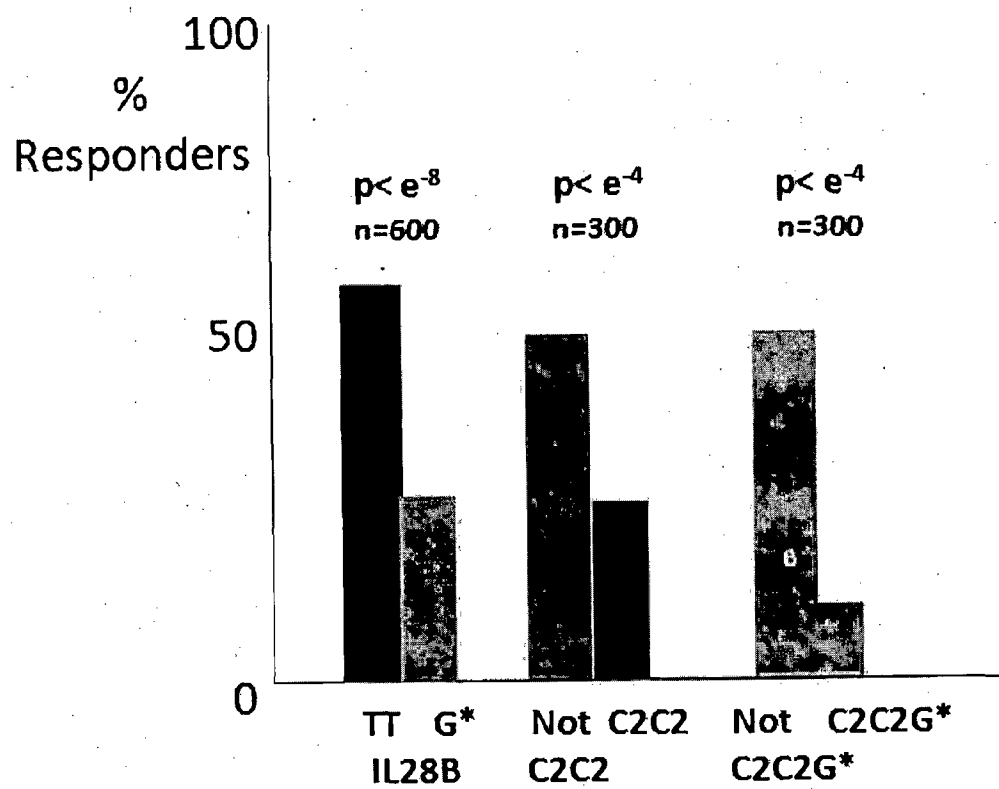


Figure 2

3/9

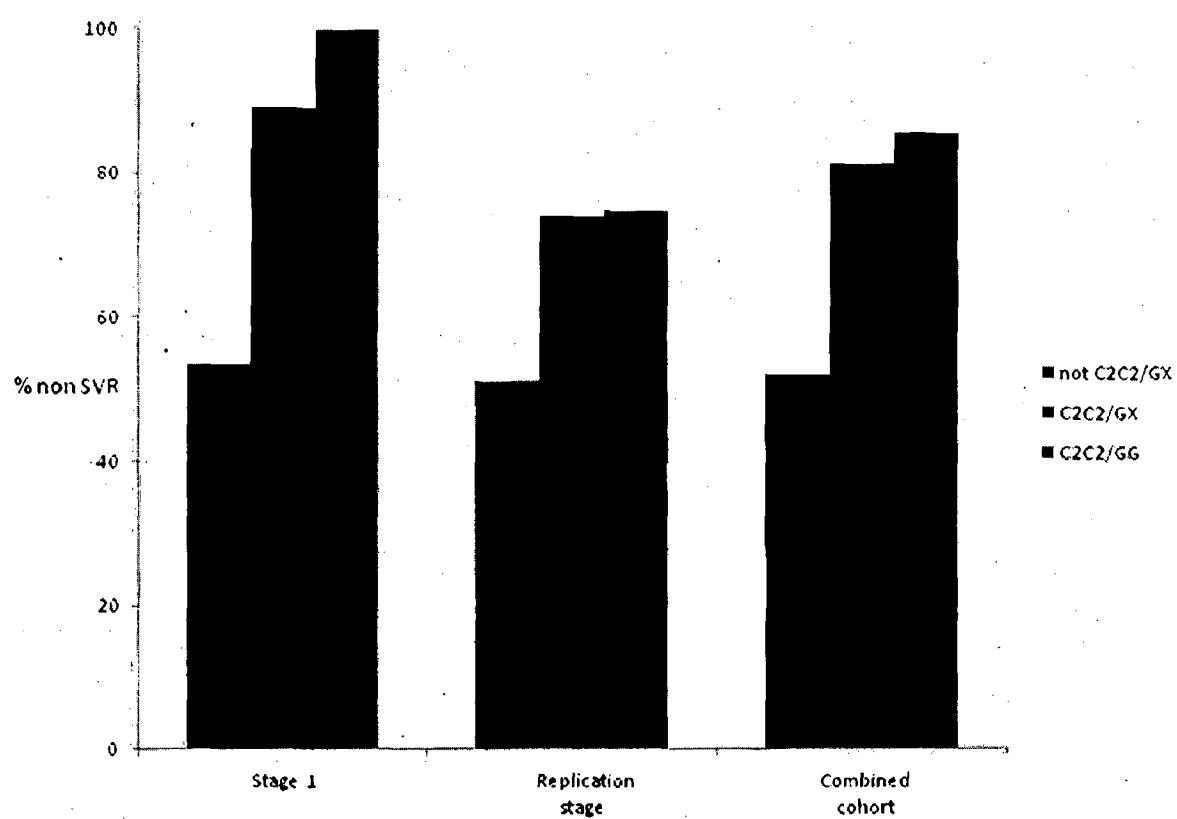


Figure 3

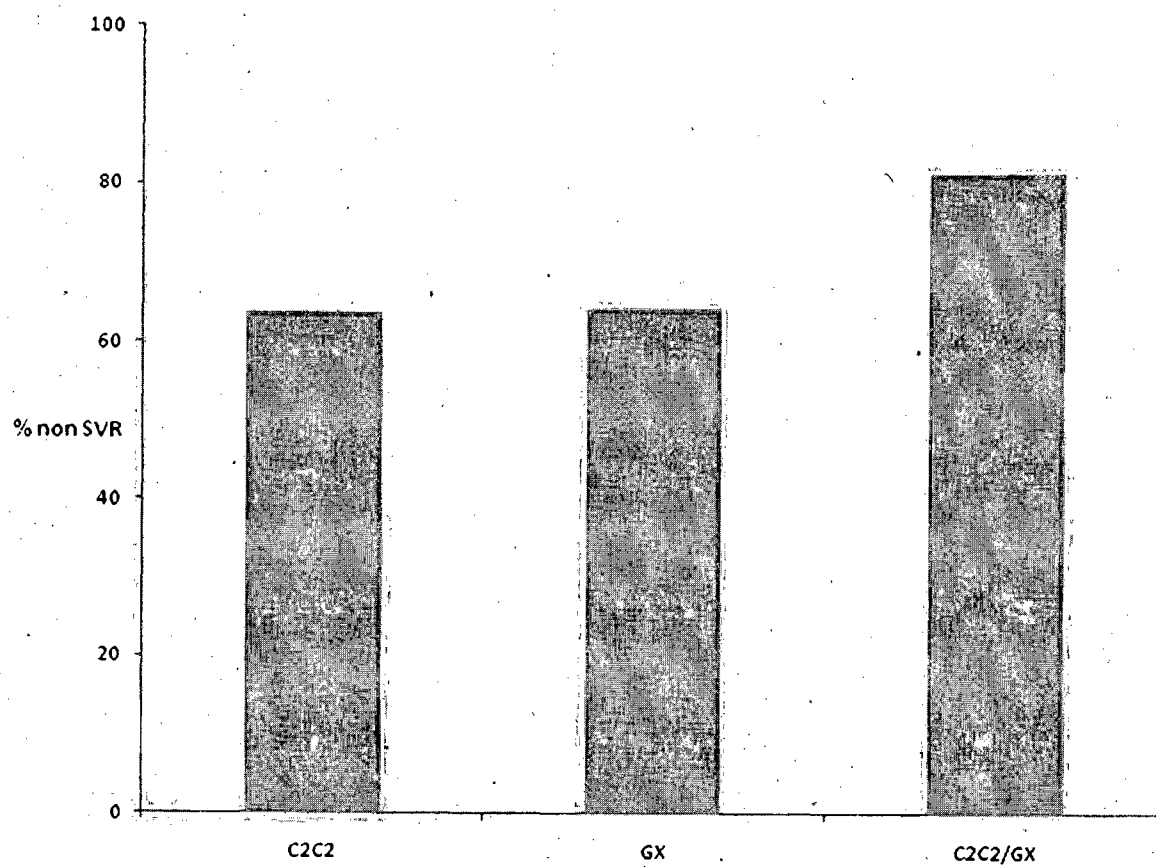


Figure 4

5/9

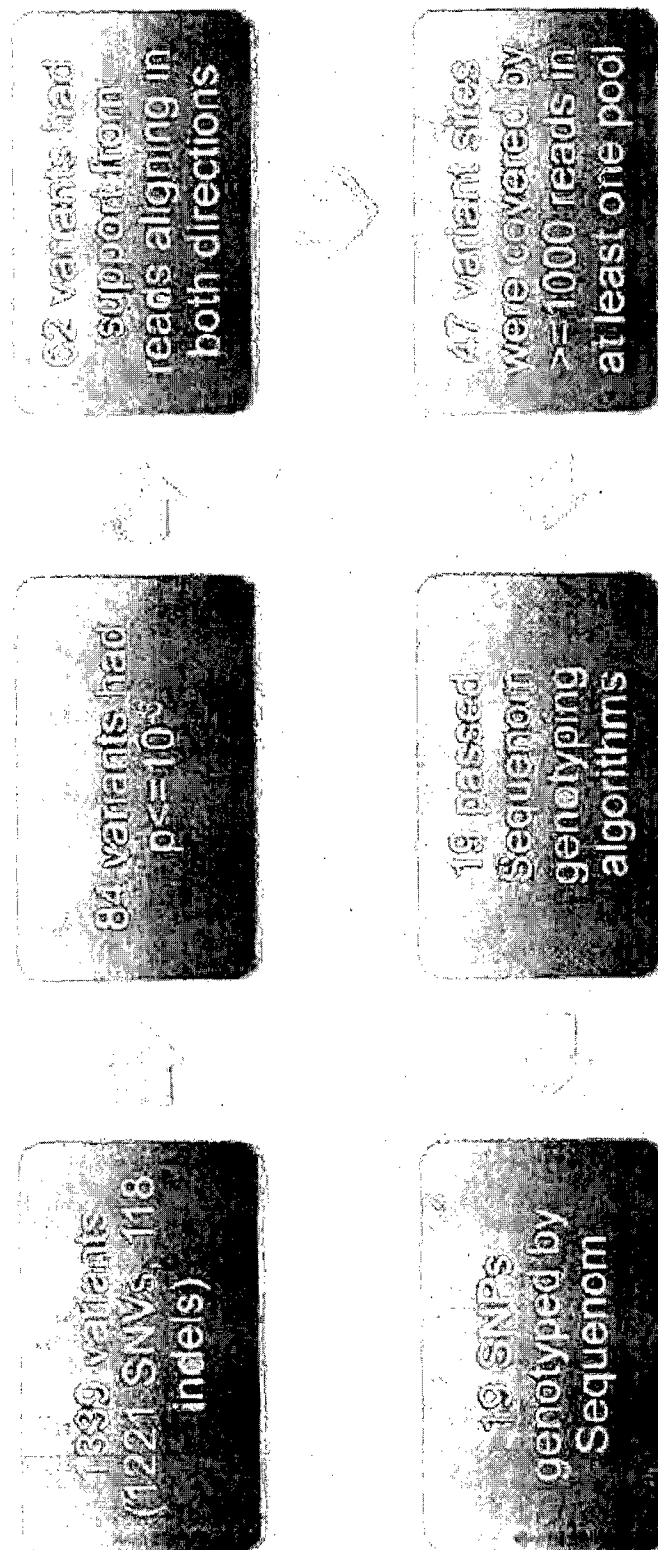


Figure 5

6/9

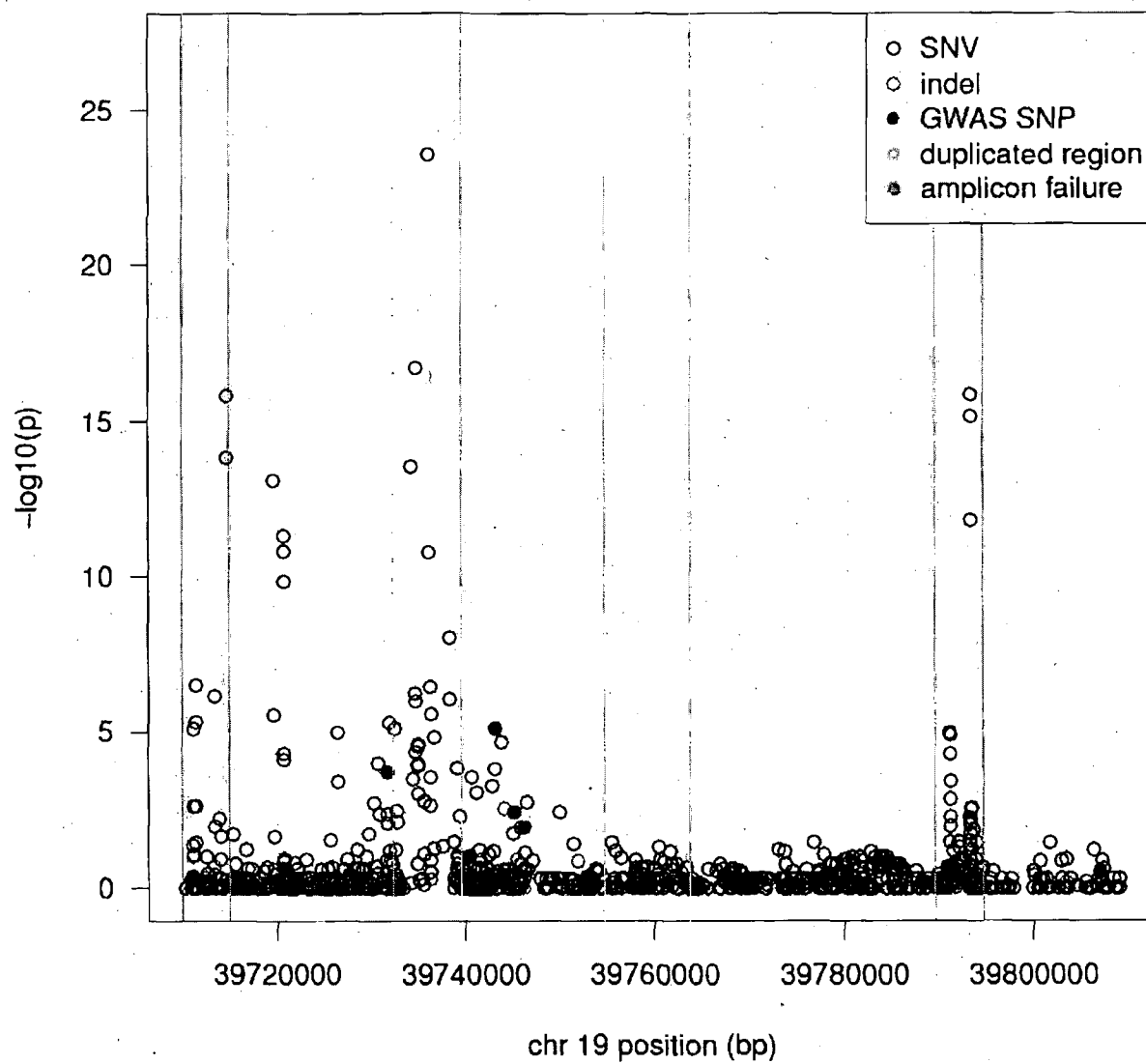


Figure 6

7/9

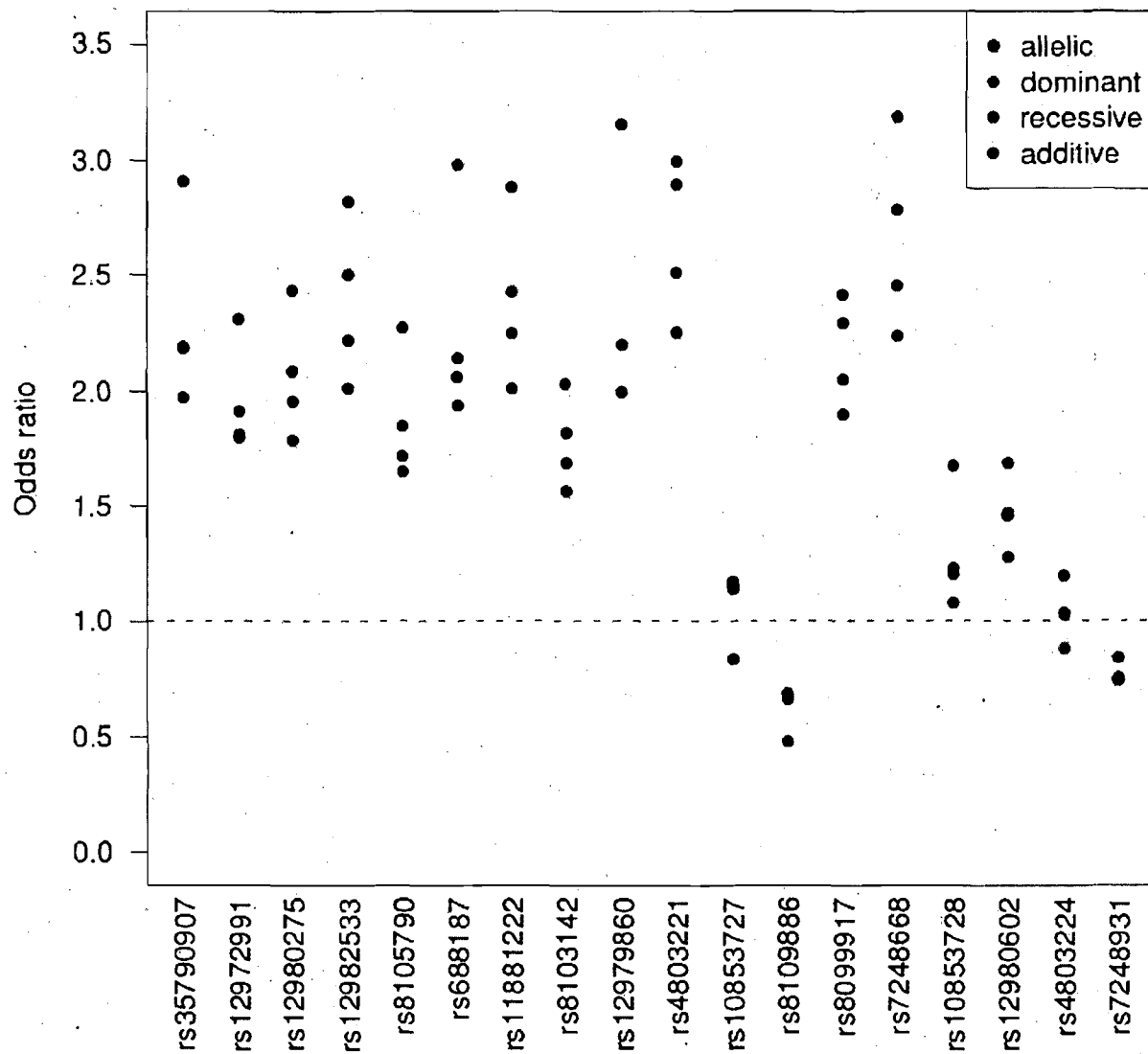


Figure 7

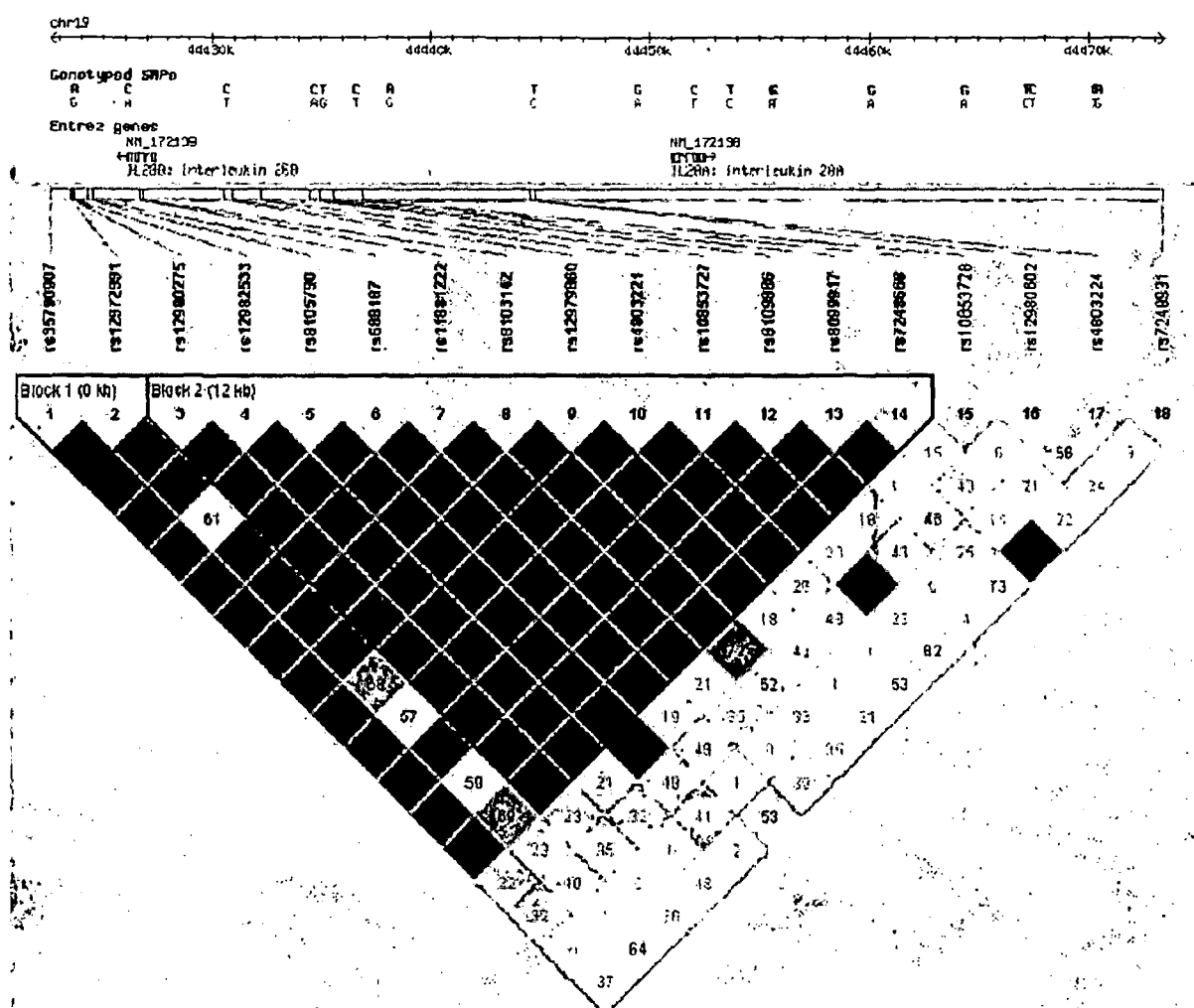


Figure 8

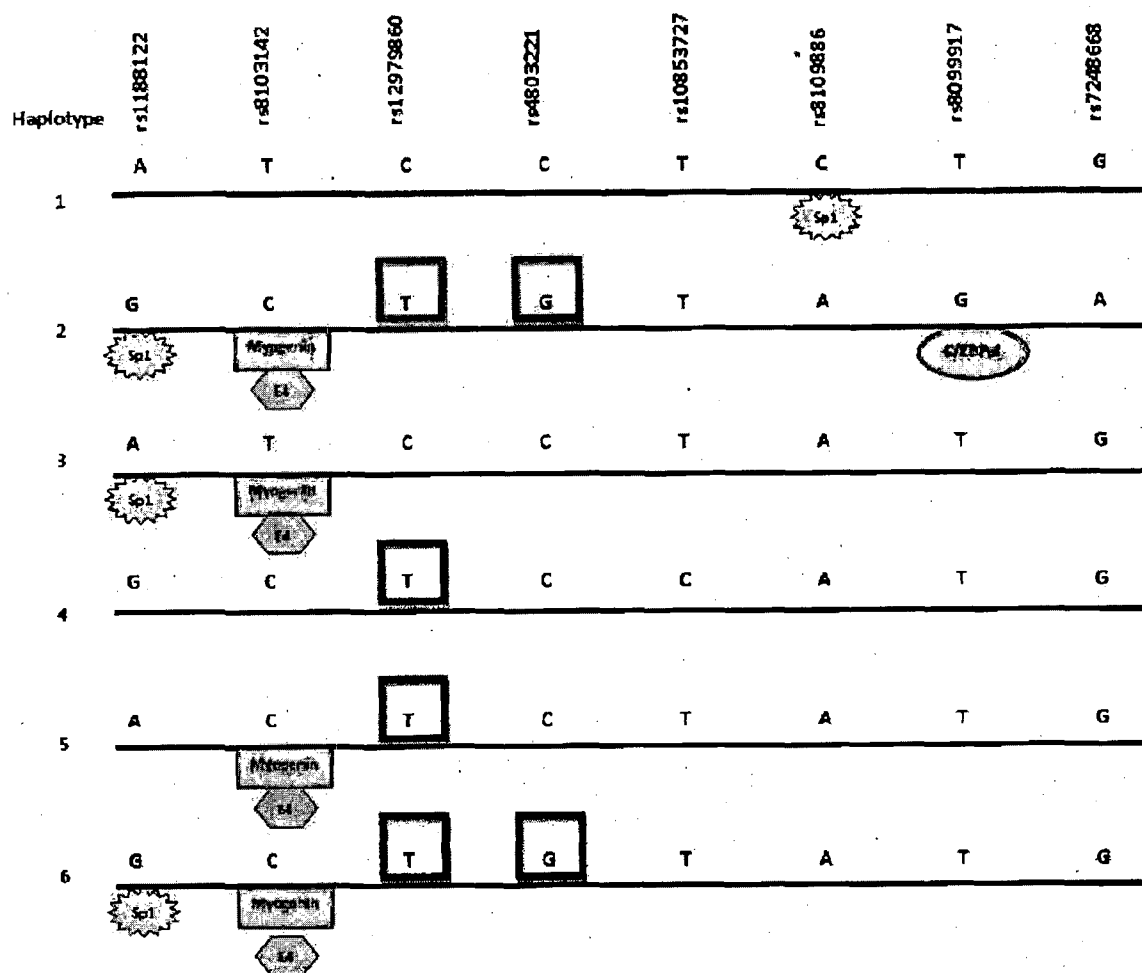


Figure 9

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/AU2012/000440

## A. CLASSIFICATION OF SUBJECT MATTER

Int. Cl.

C12Q 1/68 (2006.01)

G01N 33/48 (2006.01)

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

WPI, EDPODOC, MEDLINE, CAPLUS, BIOSIS. Keywords rs7248668, HLA-C2 and like terms

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

| Category* | Citation of document, with indication, where appropriate, of the relevant passages                                                                                                                                                                                                                    | Relevant to claim No. |
|-----------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------|
| P,X       | SMITH, K. R. <i>et al.</i> , 'Identification of improved IL28B SNPs and haplotypes for prediction of drug response in treatment of hepatitis C using massively parallel sequencing in a cross-sectional European cohort', Genome Medicine, August 2011, vol.3, no 8, pages 1-13<br>The whole document | 1-62                  |
| P,A       | WO 2011/146985 A1 (THE UNIVERSITY OF SYDNEY ET AL.) 1 December 2011                                                                                                                                                                                                                                   |                       |
| A         | WO 2010/144946 A1 (THE UNIVERSITY OF SYDNEY ET AL.) 23 December 2010                                                                                                                                                                                                                                  |                       |



Further documents are listed in the continuation of Box C



See patent family annex

|                                                                                                                                                                         |                                                                                                                                                                                                                                                  |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| * Special categories of cited documents:                                                                                                                                |                                                                                                                                                                                                                                                  |
| "A" document defining the general state of the art which is not considered to be of particular relevance                                                                | "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention                                              |
| "E" earlier application or patent but published on or after the international filing date                                                                               | "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone                                                                     |
| "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) | "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art |
| "O" document referring to an oral disclosure, use, exhibition or other means                                                                                            | "&" document member of the same patent family                                                                                                                                                                                                    |
| "P" document published prior to the international filing date but later than the priority date claimed                                                                  |                                                                                                                                                                                                                                                  |

Date of the actual completion of the international search  
18 May 2012Date of mailing of the international search report  
18 May 2012Name and mailing address of the ISA/AU  
AUSTRALIAN PATENT OFFICE  
PO BOX 200, WODEN ACT 2606, AUSTRALIA  
E-mail address: pct@ipaustalia.gov.au  
Facsimile No. +61 2 6283 7999Authorized officer  
**JULIE CHRISTIE**  
AUSTRALIAN PATENT OFFICE  
(ISO 9001 Quality Certified Service)  
Telephone No : +61 2 6283 2463

## INTERNATIONAL SEARCH REPORT

International application No.  
PCT/AU2012/000440

**Box No. I Nucleotide and/or amino acid sequence(s) (Continuation of item 1.c of the first sheet)**

1. With regard to any nucleotide and/or amino acid sequence disclosed in the international application, the international search was carried out on the basis of a sequence listing filed or furnished:
  - a. (means)
    - ☒ on paper
    - ☐ in electronic form
  - b. (time)
    - ☒ in the international application as filed
    - ☐ together with the international application in electronic form
    - ☐ subsequently to this Authority for the purposes of search
2. ☐ In addition, in the case that more than one version or copy of a sequence listing has been filed or furnished, the required statements that the information in the subsequent or additional copies is identical to that in the application as filed or does not go beyond the application as filed, as appropriate, were furnished.
3. Additional comments:

A sequence listing was provided but not used for the purposes of the search

**INTERNATIONAL SEARCH REPORT**

Information on patent family members

International application No.

**PCT/AU2012/000440**

This Annex lists the known "A" publication level patent family members relating to the patent documents cited in the above-mentioned international search report. The Australian Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

| Patent Document Cited in<br>Search Report                                                                                 |            | Patent Family Member |         |    |                      |
|---------------------------------------------------------------------------------------------------------------------------|------------|----------------------|---------|----|----------------------|
| WO                                                                                                                        | 2011146985 | NONE                 |         |    |                      |
| WO                                                                                                                        | 2010144946 | EP                   | 2443257 | SG | 176818 US 2010316608 |
| Due to data-integration issues this family listing may not include 10 digit Australian applications filed since May 2001. |            |                      |         |    |                      |
| END OF ANNEX                                                                                                              |            |                      |         |    |                      |