The definitive version of this article is published by Elsevier as:
Donaldson PT, et al Cytotoxic T lymphocyte associated antigen-4 single nucleotide polymorphisms and haplotypes in primary biliary cirrhosis. *Clinical Gastroenterology and Hepatology* 2007, 5(6), 755 - 760.
doi:10.1016/j.cgh.2007.02.012

Cytotoxic T lymphocyte associated antigen-4 single nucleotide polymorphisms and haplotypes in primary biliary cirrhosis

Short title: CTLA4 genetics in PBC

Peter Donaldson,[1,2] Sivakumar Veeramani.[1] Anna Baragiotta,[2] Annarosa Floreani,[2]
Carla Venturi,[2] Simon Pearce,[1] Valerie Wilson,[1] David Jones,[1,2]
Oliver James,[1,2] John Taylor,[3]
Julia Newton, [1,2] Margaret Bassendine.[1,2]

[1]Institute of Cellular Medicine
[2]School of Clinical Medical Sciences
[3]School of Dental Sciences
The Faculty of Medical Sciences
Newcastle University
Framlington Place
Newcastle-upon-Tyne
NE2 4HH
UK

[2]Department of Surgical and Gastroenterological Sciences,
The Medical School, University of Padova
Padova
Italy

Corresponding author

Dr Peter Donaldson
Centre for Liver Research
Institute of Cellular Medicine
4th Floor William leech Building
School of Clinical Medical Sciences
Faculty of Medical Sciences
Newcastle University
NE2 4HH
UK

Tel: 0191-222-8868
Fax: 0191 222 0723
E-mail:p.t.donaldson@ncl.ac.uk

Abbreviations: cytotoxic T lymphocyte associated antigen-4 (CTLA4): primary biliary cirrhosis (PBC): human leucocyte antigen (HLA): single nucleotide polymorphism (SNP): type 1 diabetes (T1D): quality of life (QOL): base pair (bp): haplotype tagged (ht).

Abstract

Background and Aims: Twin and family studies suggest that there is a significant heritable component to primary biliary cirrhosis (PBC). Selected cytotoxic T lymphocyte associated antigen-4 (CTLA4*)* gene polymorphisms have been proposed as "non-specific determinants of disease risk" in a variety of autoimmune diseases, including PBC. However, there has been considerable debate over the validity of these associations and the precise location of the disease promoting polymorphism. **Methods:** We have investigated six single nucleotide polymorphisms in the *CTLA4* gene in a total of 327 PBC patients and 391 healthy controls: 247 patients and 292 controls from the United Kingdom and a further 80 patients and 99 controls from northern Italy. **Results:** The previously reported association with *CTLA4 A+49G* was not replicated in the Italian series, and there were no significant differences in the distribution of any of the six polymorphisms comparing; allele, genotype or haplotype distribution in patients versus healthy controls in the UK series. Furthermore, there were no significant associations with the clinical variables; histological stage, portal hypertension or Mayo score. However, when PBC-40 Fatigue Domain scores were considered a number of significant trends were noted, but none were significant after correction for multiple testing. Thus, fatigues score were higher in those with the *CTLA4* -319 T allele ($p<0.05$, pc not significant) and in those with the *CTLA4* +49

AA genotype ($p<0.05$, pc not significant). **Conclusions:** Contrary to previous reports the *CTLA4* gene is not a major risk factor for PBC, nor is it major determinant of disease progression.

Introduction

Primary biliary cirrhosis (PBC) is an autoimmune liver disease characterised by immune mediated damage to the biliary epithelial cells lining the small intra-hepatic bile ducts[1]. Although PBC is not a Mendelian autosomal or sex-linked genetic disease, clustering of the disease in families[2-5] and specific geographic regions[6], together with the recent reports of a 63% concordance rate in monozygotic twins[7], and a sibling relative risk [λ_s] of 10.5 for PBC[5], suggest that there may be a significant heritable component to this disease.[8-10] Despite this the role of genes in PBC pathogenesis remains poorly understood and though female sex and specific HLA alleles and haplotypes[8-11] have been shown to be important determinants of disease risk, studies of other genes have produced mixed results.

Over the past 5 years there has been considerable interest in the role played by non-MHC immuno-regulatory genes in autoimmune disease, especially the T-cell regulatory gene *CTLA4* on chromosome 2q33. In particular, *CTLA4* polymorphisms have been proposed as "non-specific determinants of disease risk" in a variety of autoimmune diseases.[12] Studies of *CTLA4* polymorphism in type I diabetes and Graves' disease abound[12-14], yet there is no general consensus regarding the overall role of *CTLA4* polymorphism in autoimmunity nor is there a consensus regarding the precise location or identity of the disease causing mutations or haplotypes.

Liver disease is no exception to this pattern. Early studies reported significant associations between the *CTLA4 A+49G* SNP and the autoimmune liver diseases; PBC[15] and type 1 autoimmune hepatitis.[16] However, with the exception of an isolated report of a similar association from China,[17] these associations have not been widely replicated and recent studies have excluded the possibility of a major genetic association with the third "autoimmune" liver disease - primary sclerosing

cholangitis.[18] Disease heterogeneity which often results in case ascertainment bias, and/or the limited extent of the investigations so far may be significant factors in this failure.[19]

In PBC "disease heterogeneity" embraces such diverse phenotypic considerations as immune status, various different markers of disease progression and none-stage associated symptomatic manifestations of the disease. Of the latter, fatigue which is present in up to 50% of patients appears to have the greatest impact on patient quality of life (QOL). Despite its clinical importance, the pathogenesis of fatigue is unclear though studies have consistently shown that the severity of fatigue is independent of all other parameters of liver disease severity.[20-25]

The alternative explanation for the lack of a consensus regarding the role of *CTLA4* gene polymorphisms in autoimmune disease may be that we have been looking at the wrong region of the genome and/or overlooking significant genetic detail. The most recent studies in diabetes and Graves' disease have suggested that the key SNP within the *CTLA4* gene is not *A+49G*, but is instead an A/G SNP in the 6.1-kb region 3' of *CTLA4*.[26] This SNP (referred to as CT60) is associated with inherited variation in the efficiency of the splicing and production of soluble (s) compared with full length (fl) isoforms of CTLA-4 mRNA. Furthermore the disease susceptibility allele *CT60*G* is associated with lower levels of sCTLA-4 mRNA production. CTLA-4 is expressed on CD25 T cells and acts to down-regulate the immune response through competition with CD28 bearing T cells via CD80/CD86. It has been proposed that lower levels of sCTLA-4 in serum could lead to reduced efficiency of blocking of CD80/CD86 permitting increased or prolonged activation of CD28 T-cells.[26]

The aims of the present study were three-fold. First; to attempt to replicate our earlier findings of an association with *CTLA A+49G*[15] in a second population of PBC

patients. Second; to extend the original observations to include a more thorough analysis of *CTLA4* haplotypes in PBC, including; the 5 SNPs required for identification of the most common (tagged haplotypes) identified by Johnson *et al.*,[27] and the more recently identified diabetes-associated CT60 SNP identified by Udea *et al.*,[26]. Third; to consider in the analysis the relationship between CTLA4 polymorphism and clinical phenotype as determined by: histology score, presence or absence of portal hypertension, Mayo risk score and the symptom of fatigue.

Patients and Methods

A total of 327 well-characterized PBC patients and 391 "healthy" controls were studied. Patients and controls were recruited from two different centers as follows.

Subjects: Italian Series

80 consecutive well-characterised PBC patients with a median age of 43 years (range 18 – 73 years) were studied. All were of European ancestry and resident within the Padova area of northern Italy. Eight were male (10%) and 72 female (90%). All subjects had definite disease using standard criteria (all three of: a) liver histology diagnostic of, or compatible with, PBC, b) cholestatic liver function tests and c) positive serum anti-mitochondrial antibody titre ≥ 1:40 detected by immunofluorescence using composite tissue blocks. Subjects were excluded from the study if their biopsy (or any other clinical data) suggested additional, potentially confounding causes for liver pathology. In 76 patients liver biopsies were reviewed for confirmation of diagnosis to determine stage. Patients were classified as "advanced (late) disease" *i.e.* Scheuer stage III or IV and "early disease" *i.e.* Scheuer stage I or II.[28] Forty-one (54%) had histologically advanced disease on their last liver biopsy.

For comparison 99 Italian controls matched by geography, sex and racial origin were also studied. All subjects were of northern European Caucasoid origin and resident in and around Padova, northern Italy. All subjects and controls gave informed consent and the study was cleared by the local hospital ethics committee. DNA samples were labelled and stored by code only and analysed without prior knowledge of individual identities.

Subjects: UK Series

The UK group comprised 247 well-characterised patients with a median age of 64 years (range 23 – 85 years). All were of northern European ancestry and resident within the Newcastle area defined by postal code.[6] Twenty-seven were male (11%) and 220 female (89%). All subjects had definite disease using the same diagnostic criteria as for the Italian series. Liver histology data from within 6 months of the collection of DNA was available in 217 patients One hundred and forty-four patients (66%) had histologically advanced disease (Scheuer stage III or IV) and seventy-three (34%) had early disease on their last liver biopsy.[28] Additional data were available regarding:

1) Portal hypertension in 210 patients (defined as the presence of oesophageal varices at endoscopy and/or ascites either clinically or radiologically). By these criteria 93 patients were diagnosed as having portal hypertension
2) Mayo risk score, available in 141 patients (median 4.15; range 2.42 – 8.27)
3) QOL in 100 PBC patients assessed using the PBC-40 a fully validated disease-specific QOL measure containing a fatigue-specific domain optimised for the assessment of fatigue in PBC patients.[29]

For comparison 292 geographically and racially matched controls from north east England were studied. All subjects and controls gave informed consent and the study was cleared by the ethical committee of the Newcastle Hospitals Trust. Samples were labelled and stored by code only and analysed without prior knowledge of individual identities.

Genomic DNA was extracted from 10ml of EDTA whole blood by standard phenol/chloroform extraction protocol.

Determination of CTLA4 SNPs

For each of the SNPs tested (with the exception of CT60): 1µl of genomic DNA (20 - 200ng) was amplified in a 25µl reaction mixture containing 200µM each of dATP, dCTP, dGTP and dTTP (ABgene Surrey, UK), 1.5mM $MgCl_2$, 10mM tris-HCl pH8.3, 50mM KCl, 0.01%w/v gelatin, 0.5µM of each primer (Oswel, Southampton, UK: table 1) and 1 to 1.5U *Taq Polymerase* (ABgene) on a Perkin-Elmer GeneAmp 9700 according to the following protocols:

T-1722C

A 511bp fragment of the *CTLA4* promoter was amplified in a 25µl reaction mixture (above) according to a modification of the method of Johnson *et al.*[27] Conditions for amplification were as follows: 94^0C for 5 minutes; 35 cycles of: 94^0C for 60 seconds, 56^0C for 60 seconds, 72^0C for 60 seconds followed by a single final extension at 72^0C for 10 minutes. Following amplification, 15µl of the amplicon were digested for 3 hours with an excess of the restriction endonuclease *BbvI* (New England Biolabs (UK) Ltd, Herts, UK) at 37^0C. Digested restriction fragments were visualised on a 2% w/v agarose gels with appropriate commercially available size markers (New England Biolabs). The exchange of C for T at position -1722 creates the *BbvI* restriction site and digestion of the amplicon results in fragments of 91 and 420bp in those with the C allele.

A-1661G

A 511bp fragment of the *CTLA4* promoter was amplified as above with the same conditions for amplification. Following amplification, 15µl of the amplicon were digested for 3 hours with an excess of the restriction endonuclease *DraI* (New

England Biolabs) at 37^0C. Digested restriction fragments were visualised on a 2% w/v agarose gels with appropriate commercially available size markers (New England Biolabs). The exchange of G for A at position -1661 creates the *DraI* restriction site and digestion of the amplicon results in fragments of 70 and 441bp in those with the A allele.

C-658T

A 624bp fragment of the *CTLA4* promoter was amplified as above. Conditions for amplification were as follows: 94^0C for 5 minutes, 33 cycles of: 94^0C for 60 seconds, 56^0C for 60 seconds, 72^0C for 60 seconds followed by a single final extension at 72^0C for 10 minutes. Following amplification, 15µl of the amplicon were digested for 3 hours with an excess of the restriction endonuclease *AciI* (New England Biolabs) at 37^0C. Digested restriction fragments were visualised on a 2% w/v agarose gels with appropriate commercially available size markers (New England Biolabs). The exchange of T for C at position -658 creates the *AciI* restriction site and digestion of the amplicon results in fragments of 153 and 471bp in those with the C allele.

C-319T

A 246bp fragment of the *CTLA4* promoter was amplified as above. Conditions for amplification were as follows: 94^0C for 5 minutes, 35 cycles of: 94^0C for 60 seconds, 60^0C for 60 seconds, 72^0C for 60 seconds followed by a single final extension at 72^0C for 10 minutes. Following amplification, 10µl of the amplicon were digested for 3 hours with an excess of the restriction endonuclease *MseI* (New England Biolabs) at 37^0C. Digested restriction fragments were visualised on a 2% w/v agarose gels with appropriate commercially available size markers (New England Biolabs). The 246bp amplicon contains a constitutive *MseI* restriction site. Consequently upon digestion with *MseI* all of the amplicons were reduced in length

by 21bp to 225bp. However, the exchange of C for T at position -319 creates a further *MseI* restriction site and digestion of the amplicon results in fragments of 93 and 132bp in those with the T allele.

A+49G

A 328bp fragment of the *CTLA4* promoter was amplified in a 25µl reaction mixture (above) according to the method of Agarwal *et al.*[15] Conditions for amplification were as follows: 94^0C for 5 minutes, 35 cycles of: 94^0C for 60 seconds, 57.5^0C for 60 seconds, 72^0C for 60 seconds followed by a single final extension at 72^0C for 10 minutes. Following amplification, 15µl of the amplicon were digested for 3 hours with an excess of the restriction endonuclease *BbvI* (New England Biolabs) at 37^0C. Digested restriction fragments were visualised on a 2% w/v agarose gels with appropriate commercially available size markers (New England Biolabs). The exchange of A for G at position +49 creates the *BbvI* restriction site and digestion of the amplicon results in fragments of 84 and 244bp in those with the G allele.

CT60

A 166bp fragment of the 3' UTR of *CTLA4* was amplified in a 23.5µl reaction mixture which contained 20µl of ReddyMix™ (75mM Tris-HCl, 20mM $(NH_4)_2SO_4$, 0.01% v/v Tween 20, 0.2mM dNTPS, 1.5mM $MgCl_2$, 0.5U Taq polymerase: ABGene, Surrey, UK), 1 mM of each primer and 1µl of genomic DNA (20 - 200ng). Amplification conditions were as follows: 95^0C for 4 minutes; 35 cycles of: 95^0C for 60 seconds, 59^0C for 60 seconds, 72^0C for 60 seconds; followed by a single final extension at 72^0C for 5 minutes. 15µl of the PCR product were digested with 10 units of *HpyCH4IV* (New England Biolabs) at 37^0C and fragments visualised on 2% w/v agarose gels. The exchange of A for G creates the *HpyCH4IV* restriction site and

digestion of the CT60 amplicon results in fragments of 89 and 79bp in those with the G allele.

Data analysis and Statistics

For haplotype analysis we used the haplotype tagging (ht) method proposed by Johnson *et al.*,[27] which was developed to "significantly reduce the genotyping effort when looking for genetic determinants of common disease." One additional advantage of using this method for investigation of *CTLA4* is that the ht-SNPs have already been identified.[27] However, this information pre-dates the identification of the CT60 SNP by Udea *et al.*,[26] Therefore haplotype data were determined in two stages. First, haplotypes of the 5 tagged SNPs in the *CTLA4* gene (i.e. -1722, -1661, -658, -319, +49) were assigned manually based on the expected patterns of linkage disequilibrium.[27] These data were compared with the published data[27] and then the data for the CT60 SNP were added and the six SNP haplotypes determined based on the known patterns of linkage disequilibrium.

The genotype distribution in each sample set was tested for deviation from Hardy-Weinberg Equilibrium.

To tests for genetic associations with overall risk of PBC the distribution of genotypes and alleles and haplotypes were compared using χ^2 test and/or Fisher's exact probability test as appropriate with Statcalc on EpiInfo software (CDC Atlanta, Georgia).

To test for genetic associations within disease phenotype the relationship between genotypes, alleles and haplotypes and clinical variables was tested using χ^2 test and/or Fisher's exact probability test for the discontinuous variables and students T test for continuous variables. In this exercise presence or absence of portal

hypertension, stage I/II versus stage III/IV histology and Mayo risk scores were treated as discontinuous variables and the PBC-40 fatigue domain scores were treated as continuous variables. As there were no *a priori* associations within any of the four clinical sub-groups (histological stage, Mayo score, portal hypertension and fatigue) a correction factor of 4 was applied to all probability values generated in the analysis of phenotypes.

Results

Hardy-Weinberg Equilibrium: There were no significant deviations from the expected genotype distribution for any of the six SNP genotypes in any of the four study groups tested.

Comparison of *CTLA4* A49G SNP allele and genotype distribution in Italian PBC patients versus healthy Italian controls (table 2): There was no significant difference in the distribution of the *CTLA4 A+49G* alleles or genotypes in Italian patients versus Italian controls. Indeed the only trend seen in the Italian group was a slight increase in the G allele frequency in controls compared to patients. This was due to a small increase in the GG genotype compared with the AG genotype in controls versus patients (table 2).

Comparison of *CTLA4* SNP allele, genotype and haplotype distributions in UK PBC patients versus Healthy UK controls (tables 3, 4, 5 and 6): In the UK series there were no significant differences in the distributions of any of the six *CTLA4* SNP alleles or genotypes in patients compared to healthy controls. In addition there were no significant differences in the distribution of the six common *CTLA4* htSNP haplotypes (table 5) (as identified by the haplotype tagging method)[27] or the six-SNP extended haplotypes (table 6).

Furthermore, observed patterns of linkage disequilibrium and haplotype frequencies for the 5 tagged SNPs, were not significantly different from published patterns (see footnote to table 5).[27] There were slightly lower frequencies of haplotype 1 and slightly higher frequency of haplotype 2 in controls compared to the data of Johnson *et al.*,[27] These small variations can be attributed to population differences

between the two studies. There was also no deviation from the expected pattern of linkage disequilibrium between A49G and CT60[26] with two of the four common haplotype groups co-segregating with the *CT60*A* SNP and four preferentially co-segregating with the *CT60*G* SNP (table 6).

Comparison of *CTLA4* SNP allele, genotype and haplotype distributions in UK PBC patients by clinical subgroups: In the UK series there were no significant differences in the distribution of the *CTLA4* alleles, genotypes or haplotypes in patients with early versus late stage histology; in those with and without portal hypertension and, in those with high or low Mayo Clinic scores **(data not shown)**. However, when the PBC-40 Fatigue Domain scores were considered, significant differences were detected in the *CTLA4*-319 and the *CTLA4*+49 SNP distributions before correction for multiple testing. Thus, fatigue scores were significantly higher in those with the *CTLA4*-319 CT genotype ($p=0.04$) compared to those with the homozygote CC genotype, and possession of the CTLA4-319 T allele was associated with a significantly greater level of fatigue ($p=0.05$). Patients with *CTLA4* +49 homozygote AA genotype were also significantly more fatigued compared to those with the AG or GG genotypes ($p<0.05$). However, none of these differences were significant after correction for multiple testing.

Discussion

The present study suggests that *CTLA4* gene polymorphism does not play a major role in determining susceptibility to PBC. In addition, the absence of any clear relationship between *CTLA4* gene polymorphism and the clinical indices; histological stage, portal hypertension and Mayo Clinic score all suggest that these SNPs, alleles, genotypes and haplotypes are not major determinants of disease progression in PBC. However, there may be a relationship between the *CTLA4* gene polymorphisms and fatigue.

The opening observations of the present study are contrary to previously published data from our centre, wherein a moderate genetic association with PBC was reported.[15] Such contradictions are not unusual or without precedent. Earlier studies in type 1 diabetes (T1D), autoimmune thyroid disease and celiac sprue all reported strong to moderate effects of various *CTLA4* gene polymorphisms on disease risk.[12] Studies of the *CTLA A+49G* SNP were most successful and this SNP, which encodes a threonine to alanine substitution at position 17 in the first exon of *CTLA4* protein, became the focal point for studies of *CTLA4* in many different autoimmune diseases.[12] However, more recent studies have revealed that not only is this SNP unlikely to be a disease causing SNP in T1D and Graves' disease, but also the risk associated with the disease causing SNP is much smaller than originally suggested.[26]

There are many possible and well rehearsed explanations for the lack of replication of the findings of case control association studies.[19] In addition to the usual statistical considerations[19], we need to consider: case ascertainment bias, genotyping errors and control selection. Case ascertainment bias occurs most often when the study population comes from a tertiary referral centre. In such cases the referred patients may represent a particular clinical subgroup with less easily managed or more

severe disease. For example, one of the most recent studies of T1D has suggested that the modest increase in disease risk associated with the *CTLA4 CT60*G* allele (OR = 1.15),[26] may be confined to patients with concurrent autoimmune thyroid disease.[30] Case ascertainment bias has been reported to effect genetic associations in PBC[31], but this does not account for the current difference.

Genotyping errors occur in most studies, though it is generally thought they have minimal impact where there is adequate quality control. In the present case genotyping errors are not indicated as a major contributing factor for several reasons. First, there was a very high degree of concordance when patients from our series were independently retested; second, the genotype frequencies for the patient group presented herein are very similar to the earlier published values;[15] third, a retrospective review of the two studies indicates very few potential errors; fourth the assigned genotypes conform to the expected patterns of linkage disequilibrium.[27]

Case ascertainment bias and genotyping errors are relatively easy to identify in genetic studies and eliminate, but it is harder to identify problems related to control sample selection. Despite this the control group is as important as the patient group and a survey of major UK published series reveals similar changes in the frequency of particular SNPs in control series over time.[26, 32] In the case of *CTLA4 A+49G* there has been a general reduction in the reported frequency of the A allele in healthy controls and a compensatory increase in the G allele frequency over the past few years.[26, 32] The major difference between the published series and the present series with respect to *A+49G* is in fact related to the differences in the healthy control populations used in the two studies. The first controls group were hospital based volunteers, whilst the second were recruited through several large local businesses and may be more representative of the local population. Interestingly the current

control data are more consistent with figures from elsewhere in the United Kingdom.[26]

The present findings with respect to fatigue need to be confirmed in an independent series. As intriguing as the potential associations are they are not significant after correction for multiple testing. Given the questionable reliability of case control association studies, when they are based on small numbers[19], it is perhaps too early to draw any conclusions from this putative association. However, there are also two reasons why this unexpected result should not be completely dismissed. First, the present study includes QOL data for 100 patients only and this fatigue subgroup is not of adequate size to detect small to moderate genetic effects such as may be expected in complex disease.[19] Thus rejecting this trend out of hand may be premature and based on a type II statistical error. Second, fatigue is a debilitating problem and though studies in animal models and patients have suggested that it is central in origin and related to dysfunction of the HPA axis[33, 34] and/or altered manganese homeostasis[24] there are few firm clues as to its origins. The results of the present study, if confirmed would indicate that *CTLA4* gene polymorphisms may convey a significant risk of fatigue symptoms in PBC and this opens new avenues for future investigations offering hope to a large number (up to 50%) of PBC patients.

Overall the present study has failed to replicate earlier claims of a genetic association with *CTLA4 A+49G* and PBC[15] and it highlights, once again, the need for better planning when conducting case control association studies. Furthermore, a thorough analysis of extended haplotypes clearly demonstrates that the *CTLA4* gene does not play a major role in either disease susceptibility *per se* or disease progression in PBC. This finding is in keeping with recent observations in primary sclerosing

cholangitis[18] and the continuing controversy over the role of this gene in type 1 diabetes.[31]

References:

1. Kaplan MM, Gershwin ME. Primary biliary cirrhosis. N England J Med 2005; 353: 1261 -1273.
2. Chohan MR. Primary biliary cirrhosis in twin sisters. Gut 1973; 14: 213 - 214.
3. Bach N, Schafner F. Familial primary biliary cirrhosis. J Hepatology 1994; 20: 698 - 701.
4. Brind AM, Bray GP, Portmann BC, Williams R. Prevalence and pattern of familial disease in primary biliary cirrhosis. Gut 1995; 36: 615 - 617.
5. Jones DEJ, Watt FE, Metcalf JV, Bassendine MF, James OFW. Familial primary biliary cirrhosis reassessed: a geographically-based population study. J Hepatology 1999; 30: 402 - 407.
6. Prince MI, Chetwynd A, Diggle P, Jarner M, Metcalf JV, James OF. The geographical distribution of primary biliary cirrhosis is a well-defined cohort. Hepatology 2001; 34: 1083 - 8.
7. Selmi C, Mayo MJ, Bach N, Ishibashi H, Invernezzi P, Gish RG, Gordon SC, Wright HI, Zweiban B, Podda M. Gershwin ME. Primary biliary cirrhosis in monozygotic and dizygotic twins: genetics, epigenetics and environment. Gastroenterology 2004; 127: 485 – 492.
8. Jones DEJ, Donaldson PT. Genetic Factors in the pathogenesis of primary biliary cirrhosis. Clinics in Liver Disease 2003; 7: 841 – 864.
9. Donaldson PT. Recent Advances in Clinical Practice: Genetics of Liver Disease: Imunogenetics and Disease Pathogenesis. Gut 2004; 53: 599 – 608.
10. Tanaka A, Borchers AT, Ishibashi H, Ansari AA, Keen CL, Gershwin ME. Genetic and familial considerations of primary biliary cirrhosis. Am J Gastroenterology. 2001; 96: 8 - 15.

11. Donaldson PT, Baragiotta A, Heneghan MA, Floreani A, Venturi C, Underhill JA, Jones DEJ, James OFW, Bassendine MF. HLA class II alleles, genotypes, haplotypes and amino acids in primary biliary cirrhosis.: A large scale study. Hepatology 2006; 44: 667 – 674.
12. Kristiansen OP, Larsen ZM, Pocoit F. *CTLA-4* in autoimmune diseases - a general susceptibility gene to autoimmunity? Genes and immunity 2000; 1: 170 - 184.
13. Nistico L, Buzzeti R, Pritchard LE, van der Auwera BJ, Giovanni C, Bosi E, Larrad MT, Rios MS, Chow CC, Cockram CS, Jacobs K, Mijovic C, Bain SC, Barnett AH, Vandewalle CL, Schuit F, Gorus FK, Tosi R, Pozzilli P, Todd JA. The CTLA-4 region of chromosome 2q33 is linked to, and associated with, type 1 diabetes. Human Molecular Genet 1996; 5: 1075 – 1080.
14. Donner H, Rau H, Walfish PG, Braun J, Siegmund T, Finke A, Herwig J, Usadel KH, Badenhoop K. CTLA-4 alanine-17 confers genetic susceptibility to Grave's disease and to type 1 diabetes mellitus. J Clin Endocrinology Metabolism 1997; 82: 143 – 146.
15. Agarwal K, Jones DEJ, Daly AK, James OFW, Vaidya B, Pearce S, Bassendine MF. CTLA-4 gene polymorphism confers susceptibility to primary biliary cirrhosis. J Hepatology 2000; 32: 538 - 541.
16. Agarwal K, Czaja AJ, Jones DEJ, Donaldson PT. CTLA-4 polymorphisms and susceptibility to type 1 autoimmune hepatitis. Hepatology 31: 49 - 53, 2000.
17. Fan LY, Tu XQ, Cheng QB, Zhu Y, Feltens R, Pfeiffer T, Zhong RQ. Cytotoxic T lymphocyte associated antigen-4 gene polymorphisms confer susceptibility to primary biliary cirrhosis and autoimmune hepatitis in Chinese population. World J Gastroenterology 2004; 10: 3056 - 3059.

18. Wiencke K, Muri-Boberg K, Donaldson P, Harbo H, Ling V, Schrumpf E, Spurkland A. No major effect of the *CD28/CTLA4/ICOS* gene region on susceptibility to primary sclerosing cholangitis. Scandinavian Journal of Gastroenterology. 2006; 41: 586 – 591.
19. Colhoun HM, McKeigue PM, Davey SG. Problems of reporting genetic associations with complex outcomes. Lancet 2003; 361: 865 - 72.
20. Cauch-Dudek K, Abbey S, Stewart DE, Heathcote EJ. Fatigue in primary biliary cirrhosis. Gut 1998; 43: 705 - 710.
21. Huet PM, Deslauriers J, Tran A, Faucher C, Charbonneau J. Impact of fatigue on the quality of life in patients with primary biliary cirrhosis. Am J Gastroenterology 2000; 95: 760 - 767.
22. Goldblatt J, Taylor PJS, Lipman T, Prince M, Baragiotta A, Bassendine MF, James OFW, Jones DE. The true impact of fatigue in primary biliary cirrhosis: a population study. Gastroenterology 2002; 122: 1235 - 1241.
23. Poupon RE, Chretien Y, Chazouilleres O, Poupon R, Chwalow J. Quality of life in patients with primary biliary cirrhosis. Hepatology 2004; 40: 489 - 494.
24. Forton DM, Patel N, Prince M, Oatridge A, Hamilton G, Goldblatt J, Allsop JM, Hajnal JV, Thomas HC, Bassendine M, Jones DE, Taylor-Robinson SD. Fatigue and primary biliary cirrhosis: association of globus pallidus magnetisation transfer ratio measurements with fatigue severity and blood manganese levels. Gut 2004; 53: 587 – 592.
25. Prince MI, James OFW, Holland NP, Jones DEJ. Validation of a fatigue impact score in primary biliary cirrhosis: towards a standard for clinical and trial use. J Hepatology 2000; 32: 368 - 373.

26. Ueda H, Howson JM, Esposito L, Heward J, Snook H, Chamberlain G, Rainbow DB, Hunter KM, Smith AN, Di Genova G, Herr MH, Dahlman I, Payne F, Smyth D, Lowe C, Twells RC, Howlett S, Healy B, Nutland S, Rance HE, Everett V, Smink LJ, Lam AC, Cordell HJ, Walker NM, Bordin C, Hulme J, Motzo C, Cucca F, Hess JF, Metzker ML, Rogers J, Gregory S, Allahabadia A, Nithiyananthan R, Tuomilehto-Wolf E, Tuomilehto J, Bingley P, Gillespie KM, Undlien DE, Ronningen KS, Guja C, Ionescu-Tirgoviste C, Savage DA, Maxwell AP, Carson DJ, Patterson CC, Franklyn JA, Clayton DG, Peterson LB, Wicker LS, Todd JA, Gough SC. Association of the T-cell regulatory gene CTLA4 with susceptibility to autoimmune disease. Nature 2003; 423: 506-11.

27. Johnson GC, Esposito L, Barratt BJ, Smith AN, Heward J, Di Genova G, Ueda H, Cordell HJ, Eaves IA, Dudbridge F, Twells RC, Payne F, Hughes W, Nutland S, Stevens H, Carr P, Tuomilehto-Wolf E, Tuomilehto J, Gough SC, Clayton DG, Todd JA. Haplotype tagging for the identification of common disease genes. Nature Genetics 2001; 29: 233 - 237.

28. Scheuer PJ. Primary biliary cirrhosis Proc Roy Soc Med. 1967; 60: 1257 - 1261.

29. Jacoby A, Rannard A, Buck D, Bhala N, Newton JL, James OFW, Jones DEJ. Development, validation and evaluation of the PBC-40, a disease specific health related quality of life measure for primary biliary cirrhosis. Gut 2005; 54: 1622 – 1629.

30. Purohit S, Podolsky R, Collins C, Zheng W, Schatz D, Muir A, Hopkins S, Huang YH, She JX. Lack of correlation between the levels of soluble cytotoxic T-lymphocyte associated antigen-4 (CTLA-4) and the CT-60 genotypes. Journal of Autoimmune Diseases. 2005; 2: 8 – 13.

31. Donaldson P, Agarwal K, Craggs A, Craig W, Jones D, James O. HLA and interleukin-1 gene polymorphisms in primary biliary cirrhosis; associations with disease progression and disease susceptibility. Gut 2001; 48: 397 - 402.
32. Allahabadia A, Heward JM, Nithiyananthan R, Gibson SM, Reuser TTQ, Dodson PM, Franklyn JA, Gough SCL. MHC class II region, *CTLA4* gene, and opthalmopathy in patients with Grave's disease. Lancet 2001; 358: 984 – 985.
33. Swain MG, Le T. Chronic cholestasis in rats induces anhedonia and loss of social interest. Hepatology 1998; 28: 6 - 10.
34. Swain MG, Beck P, Rioux K, Le T. Augmented interleukin-1 beta-induced depression of locomotor activity in cholestatic rats. Hepatology 1998; 28: 1561 - 1565.

Table 1: PCR primer sequences:

SNP	Primers	Sequences
T-1722C	Forward	5'- TCC Tgg TTA CAT TTC TCC CTg AA - 3'
	Reverse	5'- TCA AgC gCC AAC AAg CAA T - 3'
A-1661G	Forward	5'- Agg AAg TgC CCA TTA ggT Tg - 3'
	Reverse	5'- TCA AgC gCC AAC AAg CAA T- 3'
C-658T	Forward	5'- AgT CTA TCC TTT TAT ggA Cgg C- 3'
	Reverse	5'- TTA CgA gAA Agg AAg CCg Tg - 3'
C-319T	Forward	5'- AAA TgA ATT ggA CTg GAT GGT - 3'
	Reverse	5'- TTA CgA gAA Agg AAg CCg Tg - 3'
A+49G	Forward	5'- CCA Cgg CTT CCT TTC TCg TA - 3'
	Reverse	5'- AgT CTC ACT CAC CTT TgC Ag - 3'
CT60	Forward	5'- CTT TgC ACC AgC CAT TAC CT - 3'
	Reverse	5'- gAA Agg ggA ggT gAA gAA CC - 3'

Table 2: CTLA+49G SNP in Italian PBC patients and controls genotype and allele distribution

		number and (percent)		
	Genotype	Patients	Controls	P value
A+49G	AA	40 (55)	54 (54)	ns
	AG	29 (40)	35 (35)	
	GG	4 (5)	10 (10)	
	Allele			
	A	109 (75)	143 (72)	ns
	G	37 (25)	55 (28)	

Footnote: All 80 Italian PBC patients and 99 Italian controls were studied. However a number of samples failed to amplify on more than one occasion and therefore the total number data on 73 patients were available for analysis.

Table 3: Summary of SNP genotype distributions in UK patients and controls

		number and (percent)		
SNP	Genotype	Patients	Controls	P value
T-1722C	CC	0	1 (0.3)	ns
	CT	44 (18)	40 (14)	
	TT	202 (82)	251 (86)	
A-1661G	AA	164 (67)	206 (71)	ns
	AG	78 (32)	81 (28)	
	GG	4 (1.6)	5 (2)	
C-658T	CC	180 (79)	239 (82)	ns
	CT	42 (19)	50 (17)	
	TT	5 (2)	1 (0.4)	
C-319T	CC	215 (87)	252 (87)	ns
	CT	31 (13)	36 (13)	
	TT	0	1 (0.3)	
A+49G	AA	75 (31)	106 (36)	ns
	AG	130 (53)	142 (49)	
	GG	38 (16)	43 (15)	
CT60	AA	32 (16)	57 (21)	ns
	AG	104 (53)	137 (50)	
	GG	59 (30)	82 (30)	

Footnote: All 247 PBC patients and 292 controls were tested for all six SNPs. However, a number of the samples failed to amplify on more than one occasion and therefore the total of data points varies for each SNP tested.

Table 4: Summary of SNP allele distributions in UK patients and controls

		number and (percent)		
SNP	Genotype	Patients	Controls	P value
T-1722C	C	44 (9)	42 (7)	ns
	T	448 (91)	542 (93)	
A-1661G	A	406 (83)	493 (84)	ns
	G	86 (17)	91 (16)	
C-658T	C	402 (89)	528 (91)	ns
	T	52 (12)	52 (9)	
C-319T	C	461 (94)	540 (93)	ns
	T	31 (6)	38 (7)	
A+49G	A	280 (58)	354 (61)	ns
	G	206 (42)	230 (39)	
CT60	A	168 (43)	251 (45)	ns
	G	222 (57)	301 (55)	

Table 5 htSNP haplotype distribution for common (>5%) *CTLA4* haplotypes

Haplotype	SNP Position					Number and Percent	
	-1722	-1661	-658	-319	+49	Patients	Controls
1	T	A	C	C	A	134 (28)	211 (36)
2	T	A	C	C	G	153 (32)	186 (32)
3	T	G	C	T	A	23 (5)	39 (7)
4	T	A	T	C	A	45 (9)	51 (9)
5	C	A	C	C	G	42 (9)	42 (7)
6	T	G	C	C	A	57 (12)	53 (9)
Other						32 (7)	0

Footnote: Haplotypes 1 – 6 above correspond to the 11 haplotypes presented by Johnson *et al.,* Though the haplotype tagging method is designed to recognise haplotypes with a frequency of 5% or greater Johnson *et al.,*[38] used additional SNPs within CTLA4 to identify some haplotypes that are less common. Thus in our scheme haplotype 1: corresponds to common haplotype A, G less common haplotypes G and J of Johnson *et al.*[38] (with a combined expected frequency of 39.3%). Haplotype 2 (above): corresponds to the common haplotype B & I (24.4%); haplotype 3: corresponds to the haplotype C only (8.6%); haplotype 4: corresponds to both haplotypes D and K (9.7%); haplotype 5: corresponds to the haplotype E (5.8%) and haplotype 6 corresponds to haplotypes F and H (7.5%).

Table 6: Six-SNP haplotype distribution for common (>5%) haplotypes.

Haplotype	SNP position						Number and Percent	
	-1722	-1661	-658	-319	+49	CT60	Patients	Controls
1	T	A	C	C	A	A	113 (31)	209 (36)
2	T	A	C	C	G	G	115 (32)	183 (32)
3	T	G	C	T	A	G	18 (5)	40 (7)
4	T	A	T	C	A	A	30 (8)	50 (9)
5	C	A	C	C	G	G	32 (9)	41 (7)
6	T	G	C	C	A	G	41 (11)	51 (9)

Footnote: Published values for this six-SNP haplotype are not available for comparison.