LETTERS

ITPA gene variants protect against anaemia in patients treated for chronic hepatitis C

Jacques Fellay[1]*, Alexander J. Thompson[2]*, Dongliang Ge[1]*, Curtis E. Gumbs[1], Thomas J. Urban[1], Kevin V. Shianna[1], Latasha D. Little[1], Ping Qiu[3], Arthur H. Bertelsen[3], Mark Watson[3], Amelia Warner[3], Andrew J. Muir[2], Clifford Brass[3], Janice Albrecht[3], Mark Sulkowski[4], John G. McHutchison[2] & David B. Goldstein[1]

Chronic infection with the hepatitis C virus (HCV) affects 170 million people worldwide and is an important cause of liver-related morbidity and mortality[1]. The standard of care therapy combines pegylated interferon (pegIFN) alpha and ribavirin (RBV), and is associated with a range of treatment-limiting adverse effects[2]. One of the most important of these is RBV-induced haemolytic anaemia, which affects most patients and is severe enough to require dose modification in up to 15% of patients. Here we show that genetic variants leading to inosine triphosphatase deficiency, a condition not thought to be clinically important, protect against haemolytic anaemia in hepatitis-C-infected patients receiving RBV.

Using DNA from consenting participants of the IDEAL study[2], we performed a genome-wide association study (GWAS) of determinants of treatment-related anaemia in individuals with chronic genotype 1 hepatitis C. A total of 1,602 DNA samples were genotyped in the context of a previously reported study of anti-HCV treatment response[3]. Following quality control steps (Supplementary Information I) 1,286 individuals, who were classified into three ethnic groups (988 European-Americans, 198 African-Americans, 100 Hispanics), were available for analyses (Table 1). Our primary analysis focused on the quantitative change in haemoglobin (Hb) levels from baseline to the fourth week of treatment—historically the point at which many patients begin erythropoietin treatment to stimulate red blood cell production. We tested each of 565,759 single nucleotide polymorphisms (SNPs) passing quality control measures in a linear regression model incorporating significant clinical covariates including baseline Hb levels.

Several SNPs on chromosome 20 (20p13 region) were found to be strongly associated with treatment-induced reduction in Hb at week 4 (Fig. 1), with the European-American sample showing overwhelming genome-wide significance ($P = 1.1 \times 10^{-45}$) for an association between quantitative Hb reduction and the SNP rs6051702. Associations with smaller effect sizes but in the same direction were observed in both the African-American and Hispanic samples (Table 2).

Further association signals were detected in the hexokinase 1 gene (*HK1*) on chromosome 10 ($P = 5.3 \times 10^{-7}$ for the intronic SNP rs10159477 in European-Americans; Supplementary Information II). This result is not genome-wide significant, but supported by other lines of evidence: rare *HK1* mutations cause severe haemolytic anaemia in both humans[4–6] and mice[7]; in a recent GWAS, *HK1* SNPs associated with differences in Hb concentration and haematocrit in Europeans[8].

The SNPs showing a genome-wide significant association with quantitative week-4 Hb reduction were spread over a 250 kilobase (kb) region that contains five different protein-coding genes (Fig. 1). We tested the independence of the top association signals in the European-American population using nested linear regression models, in which individual SNPs were added after inclusion of rs6051702, and found evidence for multiple independent signals of association: the most strongly associated SNPs after accounting for the contribution of rs6051702 have *P* values of 1.4×10^{-9} (rs2295547) and 2.3×10^{-4} (rs6051855). The persistence of strongly significant association in the region after accounting for the discovery variant suggests the possibility of multiple causal variants and/or rarer causal variants[9].

To identify candidate causal sites we searched the region for known functional variants. We focused first on the inosine triphosphatase

Table 1 | Clinical characteristics of the study population

	Population		
	European-Americans	African-Americans	Hispanics
n	988	198	100
Sex (F/M)	378/610	78/120	36/64
Age (yrs)	47.3 (7.4)	49.7 (6.6)	44.8 (9.3)
BMI (kg m^{-2})	27.9 (4.5)	29.7 (5.0)	29.3 (5.4)
Baseline weight (kg)	83.3 (16.1)	88.7 (14.3)	83.0 (16.7)
Baseline liver fibrosis stage* (*n*, %)			
Minimal (F0–2)	876 (88.7%)	182 (91.9%)	86 (86.0%)
Advanced (F3–4)	112 (11.3%)	16 (8.1%)	14 (14.0%)
Initial daily RBV dose† (*n*, %)			
800 mg	88 (8.9%)	4 (2.0%)	6 (6.0%)
1,000 mg	377 (38.2%)	63 (31.8%)	41 (41.0%)
1,200 mg	460 (46.6%)	117 (59.1%)	45 (45.0%)
1,400 mg	63 (6.4%)	14 (7.1%)	8 (8.0%)
Peg-interferon treatment‡			
PegIFNα-2a	330 (33.4%)	66 (33.3%)	31 (31.0%)
PegIFNα-2b 1.0	333 (33.7%)	69 (34.9%)	32 (32.0%)
PegIFNα-2b 1.5	325 (32.9%)	63 (31.8%)	37 (37.0%)
Baseline Hb value (g dl^{-1})	15.1 (1.2)	14.6 (1.2)	15.2 (1.3)
Hb reduction at week 4, mean (g dl^{-1})	2.9 (1.5)	2.4 (1.3)	3.2 (1.5)
Severe anaemia, Hb <10 g dl^{-1} at week 4 (*n*, %)	89 (9.0%)	18 (9.1%)	11 (11.0%)

Data shown are mean with s.d. in parentheses unless otherwise indicated. BMI, body mass index.

* Liver biopsies were evaluated for METAVIR fibrosis staging. The METAVIR scoring system classifies fibrosis on a 5-point scale: F0, no fibrosis; F1, portal fibrosis without septa; F2, few septa; F3, numerous bridging septa without cirrhosis; F4, cirrhosis.

† Initial daily RBV dose was weight-based on a sliding scale in subjects' baseline weight, from 40–125 kg.

‡ PegIFNα-2b 1.0 and PegIFNα-2b 1.5 refer to 1.0 μg kg^{-1} week^{-1} and 1.5 μg kg^{-1} week^{-1} interferon, respectively.

[1]Institute for Genome Sciences & Policy, Center for Human Genome Variation, Duke University, Durham, North Carolina 27708, USA. [2]Duke Clinical Research Institute and Division of Gastroenterology, School of Medicine, Division of Gastroenterology, Duke University, Durham, North Carolina 27705, USA. [3]Schering-Plough Research Institute, Kenilworth, New Jersey 07033, USA. [4]Johns Hopkins University School of Medicine, Baltimore, Maryland 21205, USA.
*These authors contributed equally to this work.

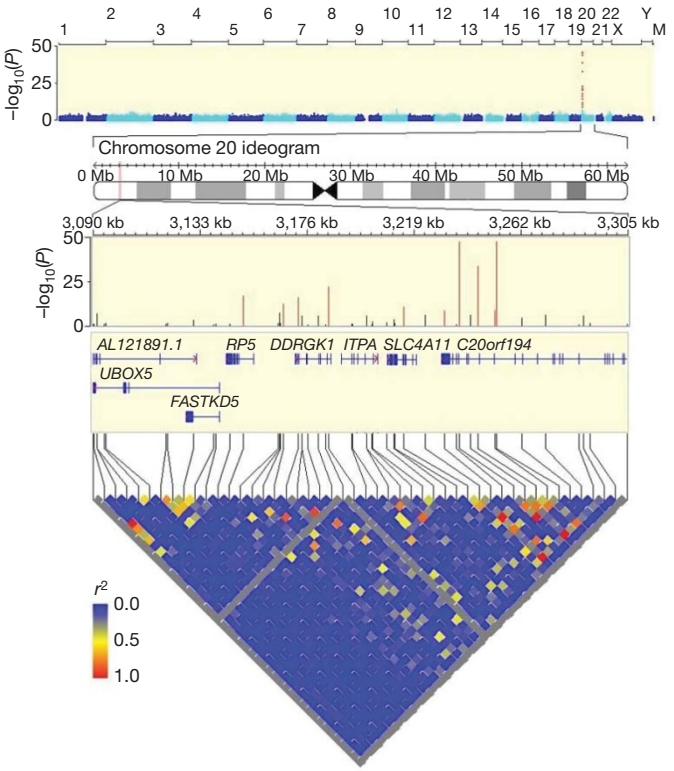

Figure 1 | Genomic overview of the 20q13 region including the genome-wide significant associated variants and the *ITPA* gene. Indicated are the *P* values [$-\log_{10}(P)$] of all genotyped SNPs in the region and the structures of the surrounding genes. The SNPs that show genome-wide significant association with quantitative reduction in haemoglobin levels are marked in red. The results were annotated using the WGAViewer software[27].

(*ITPA*) gene, which encodes a protein that hydrolyses inosine triphosphate (ITP). Several gene mutations have been described that lead to ITPA deficiency, a benign red cell enzymopathy characterized by the accumulation of ITP in erythrocytes and increased toxicity of purine analogue drugs[10,11]. In particular, reduced ITPA activity has been documented for a missense variant in exon 2 (rs1127354, resulting in a proline-to-threonine substitution denoted P32T) and a splicing-altering SNP located in the second intron (rs7270101)[12–15]. Several studies have documented and replicated the functional effects of these polymorphisms, showing that both minor alleles independently reduce ITPA activity and that homozygosity for the P32T mutation results in non-detectable ITPA activity and strong accumulation of ITP in red blood cells (Supplementary Table 1)[12,16–18]. Both SNPs can therefore be considered validated functional variants.

Using HapMap data from CEU parents (Utah residents with ancestry from northern and western Europe)[19] we found both of these polymorphisms to be associated with rs6051702, with the low-activity variants preferentially associating with the protective rs6501702 C allele. We can directly test whether the association signal reflects the functional alleles by considering them equivalent and defining a new allele as a 'low-activity allele' made up of either functional variant (Supplementary Information III). We tested this allele for association with surrounding polymorphisms and found that of all the HapMap SNPs located in the surrounding 1 megabase (Mb)

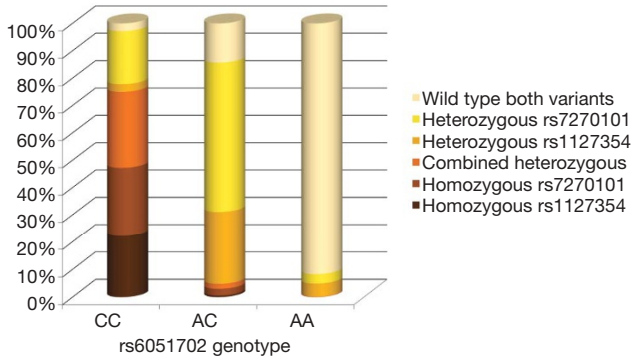

Figure 2 | Two *ITPA* polymorphisms known to be responsible for inosine triphosphatase deficiency co-segregate with the rs6051702 C allele that strongly associates with protection against Hb reduction in European-Americans. In the European-American population included in the study, the low-activity *ITPA* variants rs1127354 (missense P32T) and rs7270101 (intronic, splicing-altering) are found almost exclusively on chromosomes that also carry the rs6051702 minor allele C (Supplementary Table 4). The graph shows, for each rs6051702 genotype, the percentages of individuals that are either heterozygous or homozygous for at least one *ITPA* variant.

region, the highest r^2 is 0.65, which is what is observed for our discovery variant (Supplementary Table 2).

This observation suggested the possibility that these two known functional *ITPA* variants are responsible for part or all of the observed association and confer protection against anaemia. To test this possibility, we first sequenced the entire coding region of the gene in a subset of 168 samples, and found no other obvious reduced function mutations (Supplementary Table 3). We then genotyped the known functional SNPs rs1127354 and rs7270101 in our entire cohort. We observed that in European-Americans, the two functional *ITPA* variants were found almost exclusively on chromosomes that also carry the rs6051702 minor allele C (Fig. 2 and Supplementary Table 4). Moreover, when these two variants were incorporated into a regression model, they entirely explained the association signal initially identified (Supplementary Table 5). Finally, in European-Americans, each functional variant was strongly independently associated with protection (Table 3). A similar picture was observed in African-Americans and in Hispanics: the SNP that has the strongest association with treatment-induced Hb reduction in the region was different between populations (Table 2), but the signal is explained in each case by the associations between the most significant SNP and the *ITPA* low-activity variants. When the evidence for each functional SNP is combined across the populations, we find that the *P* values strengthen considerably to 1.7×10^{-58} for rs1127354 and 5.9×10^{-26} for rs7270101 (Table 3). Therefore, several lines of evidence confirm that the two known functional variants conferring reduced ITPA activity are responsible for the protection against anaemia identified in the original GWAS.

To assess the affect of both functional alleles on anaemia, we considered the combined low-activity allele described earlier in an additive regression model. We found that the combined allele has an association of $P = 2.2 \times 10^{-91}$ and that the resulting difference in ITPA function explains 28.7% of the variability in quantitative Hb reduction in the European-American sample, 19.1% in African-Americans and 23.2% in Hispanics.

We made a direct assessment of the clinical relevance of these variants by considering the proportion of patients suffering clinically

Table 2 | GWAS of quantitative Hb reduction at week 4

SNP	Sample in which the SNP shows the strongest association	*P* value in European-Americans ($n = 988$)	*P* value in African-American ($n = 198$)	*P* value in Hispanics ($n = 100$)
rs6051702	European-Americans	1.1×10^{-45}	1.9×10^{-1}	9.5×10^{-3}
rs3810560	African-Americans	2.6×10^{-2}	1.2×10^{-4}	3.0×10^{-1}
rs11697114	Hispanics	2.0×10^{-6}	2.8×10^{-2}	2.1×10^{-4}

The 20p13 region SNPs showing the strongest association in each of the three study populations are reported, together with their *P* values in the different ethnic groups.

Table 3 | *ITPA* variants protect against Hb reduction

ITPA variant	Population	MAF (%)	*P* value	Independent *P* value
rs1127354	European-Americans	7.6	4.6×10^{-52}	2.3×10^{-68}
	African-Americans	4.6	2.7×10^{-7}	5.1×10^{-7}
	Hispanics	4.0	1.2×10^{-3}	5.6×10^{-5}
	All (combined)	6.9	1.7×10^{-58}	5.9×10^{-26}
rs7270101	European-Americans	12.3	6.8×10^{-22}	3.6×10^{-38}
	African-Americans	7.9	3.0×10^{-5}	6.6×10^{-5}
	Hispanics	8.0	3.8×10^{-4}	1.9×10^{-5}
	All (combined)	11.2	8.5×10^{-76}	2.6×10^{-43}

The two *ITPA* functional SNPs rs1127354 and rs7270101 show strong independent association with protection against treatment-induced Hb reduction. Independent *P* values were calculated in models in which the other functional variant was already included. Combined *P* values for all three populations were obtained using the weighted *Z*-method[26]. MAF, minor allele frequency.

significant anaemia, which we defined as either a decline in Hb of $>3\,g\,dl^{-1}$ or Hb levels $<10\,g\,dl^{-1}$, which is the threshold at which RBV dose reduction is recommended. As depicted in Fig. 3, we modelled this using both the individual genotypes and the degree of ITPA deficiency, predicted from the residual enzymatic activity measured in the presence of the two functional variants[12,16–18]. We found that of 184 patients predicted to have deficient ITPA function corresponding to less than one-third of the normal enzymatic activity, none was observed to have Hb levels $<10\,g\,dl^{-1}$, and only seven (3.8%) had a decline in Hb of $>3\,g\,dl^{-1}$ at week 4. On the other hand, of the 863 patients with predicted normal ITPA function, 11.7% had Hb levels $<10\,g\,dl^{-1}$ and 55.9% had a decline in Hb of $>3\,g\,dl^{-1}$.

The mechanism of RBV-induced haemolytic anaemia remains poorly understood[20], but clearly involves an accumulation of active forms of RBV in red blood cells, including the triphosphate form (RBV-TP). Similarly, it has been well-documented that inosine triphosphatase deficiency leads to an accumulation of ITP in red blood cells[16,21]. A simple model to explain these observations would be that

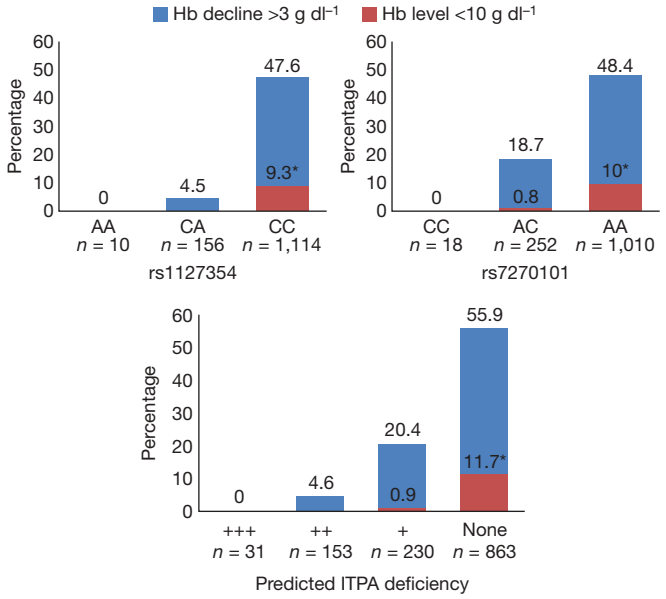

Figure 3 | ITPA deficiency protects against clinically significant decline in Hb concentration induced by HCV anti-viral treatment. Percentages of treated subjects with Hb decline of $>3\,g\,dl^{-1}$ (blue) or Hb concentrations $<10\,g\,dl^{-1}$ (red) at week 4 are shown for the two *ITPA* low function variants rs1127354 and rs7270101 (top), and for different degrees of predicted ITPA deficiency (bottom). Severity of ITPA deficiency was estimated from refs 12, 16–18: compared to wild type, ITPA activity decreased to 60% with rs7270101 heterozygosity (+); to 30% with rs1127354 heterozygosity or rs7270101 homozygosity (++); and to very low residual activity with combined heterozygosity or rs1127354 homozygosity (+++). Asterisk denotes that six subjects had Hb concentrations $<10\,g\,dl^{-1}$ at week 4 but Hb decline of $<3\,g\,dl^{-1}$.

ITP competes with RBV-TP in whatever cellular processes are affected by RBV-TP, and thereby protects cells from the lytic effects of RBV-TP. An alternative explanation is that ITPA activity and/or ITP levels directly or indirectly influence RBV pharmacokinetics. We did not, however, observe any significant association between the *ITPA*-deficiency variants and early or late anti-HCV treatment outcomes (Supplementary Information IV), which suggests that RBV pharmacokinetics is not significantly altered in hepatic cells. We also evaluated global differences in the frequencies of the *ITPA* variants. We found that although each of the two functional variants showed strong geographic differences in frequency (Supplementary Table 6), these are not correlated, and the predicted overall ITPA activity levels do not seem to show important variation. This suggests that these variants do not create strong population-level differences in susceptibility to anaemia, as was observed for the effect of *IL28B* variation on treatment response[3].

Two related features of these observations are worth emphasizing. First, the *ITPA* variants constitute a clear example of a synthetic association in which the effects of rarer functional variants are observed as an association for a more common variant present on a whole-genome genotyping chip: indeed, the minor-allele frequency is higher for the top-associated SNP rs6051702 (19.4%) than for the causal variants rs1127354 (7.6%) and rs7270101 (12.3%) in European-Americans. Second, ITPA deficiency seems to behave as a classical pharmacogenetic trait as first described by Motulsky[22], in which genetic variation that is otherwise innocuous confers a strong drug response phenotype. It is interesting to note that, in contrast with many common diseases[23], relatively common genetic variants have been identified by GWAS that explain very significant proportions of the population variation in drug responses, including the *ITPA* variants described here and the *IL28B* variants reported recently[3]. It is possible that this difference stems in part from the fact that drug responses represent new environmental challenges, and that variants with a strong effect would not necessarily be purged by selection, as would be the case for variants affecting relatively early-onset diseases. We note, however, that like many common diseases, drug responses are also likely to be heavily influenced by variants that are too rare for effective representation in GWAS[23].

Finally, because ITPA deficiency seems to be a benign condition, it may be possible to protect against RBV-induced anaemia by pharmacological intervention against ITPA. The identification of inosine triphosphatase deficiency as a major projective factor against RBV-induced haemolytic anaemia therefore not only provides a valuable pharmacogenetic diagnostic, but also a window into red blood cell biology and the processes governing lysis.

METHODS SUMMARY

All participants were included in the IDEAL study[2], which compared the effectiveness of three anti-HCV treatment regimens including RBV. The genotype data used was described previously (Supplementary Information I)[3].

Haemoglobin values were measured at baseline, at weeks 2, 4, 8 and 12, and then every 6 weeks up to treatment completion (48 weeks). Follow-up measurements were obtained 4, 12 and 24 weeks after treatment. We considered Hb decline at week 4 to be a clinically important time point at which significant anaemia had occurred but no growth factor therapy had been instituted. We excluded patients who were <80% adherent to pegIFN or RBV up to week 4 ($n=95$), and patients with missing week-4 Hb data ($n=21$). Three phenotypes were considered: (1) absolute reduction in Hb, (2) a reduction in Hb of $>3\,g\,dl^{-1}$, and (3) a reduction in Hb levels to $<10\,g\,dl^{-1}$. The threshold of Hb reduction of $>3\,g\,dl^{-1}$ was chosen as a clinically significant Hb decline; the threshold of Hb $<10\,g\,dl^{-1}$ was chosen because it is the level at which RBV dose reduction is recommended according to package insert.

Our association tests involved single-marker genotype trend tests of association between each SNP and the phenotypes, using linear and logistic regression models implemented in PLINK[24]. Covariates included age, gender, weight, fibrosis severity on pretreatment liver biopsy, baseline Hb level, RBV dose, and type/dose of pegIFN. To control for the possibility of spurious associations resulting from population stratification, we used a modified EIGENSTRAT method[25] (Supplementary Information I): we defined three ethnic groups, in

which separate analyses were run, and corrected for population ancestry within each group. Combined *P* values for all three populations were obtained using the Stouffer's weight *Z*-method[17]. We assessed significance with Bonferroni correction using the total number of tests (565,759) as the denominator for the calculation (P cutoff $= 8.8 \times 10^{-8}$).

Received 13 November 2009; accepted 11 January 2010.
Published online 21 February 2010.

1. Lavanchy, D. The global burden of hepatitis. *Liver Int.* **29,** 74–81 (2009).
2. McHutchison, J. G. *et al.* Peginterferon α-2b or α-2a with ribavirin for treatment of hepatitis C infection. *N. Engl. J. Med.* **361,** 580–593 (2009).
3. Ge, D. *et al.* Genetic variation in IL28B predicts hepatitis C treatment-induced viral clearance. *Nature* **461,** 399–401 (2009).
4. van Wijk, R. *et al.* HK Utrecht: missense mutation in the active site of human hexokinase associated with hexokinase deficiency and severe nonspherocytic hemolytic anemia. *Blood* **101,** 345–347 (2003).
5. de Vooght, K. M. K., van Solinge, W. W., van Wesel, A. C., Kersting, S. & van Wijk, R. First mutation in the red blood cell-specific promoter of hexokinase combined with a novel missense mutation causes hexokinase deficiency and mild chronic hemolysis. *Haematologica* **94,** 1203–1210 (2009).
6. Bonnefond, A. *et al.* A genetic variant in *HK1* is associated with pro-anemic state and HbA1c but not other glycemic control related traits. *Diabetes* **58,** 2687–2697 (2009).
7. Peters, L. L. *et al.* Downeast anemia (*dea*), a new mouse model of severe nonspherocytic hemolytic anemia caused by hexokinase (HK_1) deficiency. *Blood Cells Mol. Dis.* **27,** 850–860 (2001).
8. Ganesh, S. K. *et al.* Multiple loci influence erythrocyte phenotypes in the CHARGE Consortium. *Nature Genet.* **41,** 1191–1198 (2009).
9. Dickson, S. P., Wang, K., Kranz, I., Hakonarson, H. & Goldstein, D. B. Rare variants create synthetic genome-wide associations. *PLoS Biol.* **8,** e1000294 (2010).
10. Bierau, J., Lindhout, M. & Bakker, J. A. Pharmacogenetic significance of inosine triphosphatase. *Pharmacogenomics* **8,** 1221–1228 (2007).
11. Stocco, G. *et al.* Genetic polymorphism of inosine triphosphate pyrophosphatase is a determinant of mercaptopurine metabolism and toxicity during treatment for acute lymphoblastic leukemia. *Clin. Pharmacol. Ther.* **85,** 164–172 (2009).
12. Sumi, S. *et al.* Genetic basis of inosine triphosphate pyrophosphohydrolase deficiency. *Hum. Genet.* **111,** 360–367 (2002).
13. Cao, H. & Hegele, R. A. DNA polymorphisms in ITPA including basis of inosine triphosphatase deficiency. *J. Hum. Genet.* **47,** 620–622 (2002).
14. Arenas, M., Duley, J., Sumi, S., Sanderson, J. & Marinaki, A. The *ITPA* c.94C>A and g.IVS2+21A>C sequence variants contribute to missplicing of the *ITPA* gene. *Biochim. Biophys. Acta* **1772,** 96–102 (2007).
15. Stepchenkova, E. I. *et al.* Functional study of the P32T ITPA variant associated with drug sensitivity in humans. *J. Mol. Biol.* **392,** 602–613 (2009).
16. Shipkova, M., Lorenz, K., Oellerich, M., Wieland, E. & von Ahsen, N. Measurement of erythrocyte inosine triphosphate pyrophosphohydrolase (ITPA) activity by HPLC and correlation of *ITPA* genotype-phenotype in a Caucasian population. *Clin. Chem.* **52,** 240–247 (2006).
17. Atanasova, S. *et al.* Analysis of ITPA phenotype-genotype correlation in the Bulgarian population revealed a novel gene variant in exon 6. *Ther. Drug Monit.* **29,** 6–10 (2007).
18. Maeda, T. *et al.* Genetic basis of inosine triphosphate pyrophosphohydrolase deficiency in the Japanese population. *Mol. Genet. Metab.* **85,** 271–279 (2005).
19. The International HapMap Consortium. A haplotype map of the human genome. *Nature* **437,** 1299–1320 (2005).
20. Russmann, S., Grattagliano, I., Portincasa, P., Palmieri, V. O. & Palasciano, G. Ribavirin-induced anemia: mechanisms, risk factors and related targets for future research. *Curr. Med. Chem.* **13,** 3351–3357 (2006).
21. Fraser, J. H., Meyers, H., Henderson, J. F., Brox, L. W. & McCoy, E. E. Individual variation in inosine triphosphate accumulation in human erythrocytes. *Clin. Biochem.* **8,** 353–364 (1975).
22. Motulsky, A. G. Drug reactions enzymes, and biochemical genetics. *J. Am. Med. Assoc.* **165,** 835–837 (1957).
23. Goldstein, D. B. Common genetic variation and human traits. *N. Engl. J. Med.* **360,** 1696–1698 (2009).
24. Purcell, S. *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81,** 559–575 (2007).
25. Price, A. L. *et al.* Principal components analysis corrects for stratification in genome-wide association studies. *Nature Genet.* **38,** 904–909 (2006).
26. Whitlock, M. C. Combining probability from independent tests: the weighted *Z*-method is superior to Fisher's approach. *J. Evol. Biol.* **18,** 1368–1373 (2005).
27. Ge, D. *et al.* WGAViewer: software for genomic annotation of whole genome association studies. *Genome Res.* **18,** 640–643 (2008).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We are indebted to the IDEAL principal investigators, the study coordinators, nurses and patients involved in the study. We also thank E. Gustafson, P. Savino, D. Devlin, S. Noviello, M. Geffner, E. L. Heinzen, A. C. Need and E. T. Cirulli for their contributions to the study. This study was funded by the Schering-Plough Research Institute, Kenilworth, New Jersey. A.J.T. receives funding support from the National Health and Medical Research Council of Australia and the Gastroenterological Society of Australia.

Author Contributions J.F., A.J.T. and D.G. contributed equally. D.B.G., J.G.M., J.A., A.J.T. and J.F. defined the clinical phenotypes. J.F., A.J.T., D.G. and T.J.U. performed the statistical and bioinformatical analyses. K.V.S. performed the genotyping. C.E.G. and L.D.L. performed the sequencing experiments. P.Q., A.H.B., M.W., A.W., A.J.M., M.S., C.B. and J.A. collected and analysed the clinical data. A.J.M., M.S., J.G.M. and D.B.G designed the study. J.F., A.J.T. and D.B.G. wrote the manuscript with critical input from all authors.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare competing financial interests: details accompany the full-text HTML version of the paper at www.nature.com/nature. Correspondence and requests for materials should be addressed to D.B.G. (d.goldstein@duke.edu) or J.G.M. (mchut001@mc.duke.edu).