Specific type 1 diabetes risk genes underpin age-at-diagnosis and indicate joint defects in immunity, beta-cell fragility and responses to viral infections in early-onset disease.

Inshaw JRJ, Cutler AJ, Crouch DJM, Wicker LS and Todd JA

JDRF/Wellcome Diabetes and Inflammation Laboratory, Wellcome Centre for Human Genetics, University of Oxford, United Kingdom

Corresponding authors: Mr Jamie Inshaw, Professor John Todd

JDRF/Wellcome Diabetes and Inflammation Laboratory,

Wellcome Centre for Human Genetics,

NIHR Oxford Biomedical Research Centre,

Nuffield Department of Medicine,

Roosevelt Drive,

Oxford,

OX3 7BN

01865 287859

jinshaw@well.ox.ac.uk

jatodd@well.ox.ac.uk

Abstract

Background

Immunohistological analyses of pancreata from patients with autoimmune type 1 diabetes (T1D) suggest a stratification of islet pathology of both B and T lymphocyte islet inflammation common in children diagnosed under age 7 years, whereas B cells are rare in those diagnosed age ≥13. Based on these observations, we would expect to see genetic susceptibility differences between these age-at-diagnosis groups at the population level. Moreover, these genetic susceptibility differences could inform us on the aetiology of this most aggressive form of T1D that initiates in the first years of life.

Methods

Using multinomial logistic regression models we tested if the known T1D loci (17 within the human leucocyte antigen (HLA) region and 55 others, non HLA regions) had significantly stronger effect sizes in the <7 group compared to the ≥13 group, using genotype data from 26,991 individuals (18,400 controls, 3,111 T1D diagnosed <7 years of age, 3,759 at 7-13 and 1,721 at ≥13).

Findings

Six associations of the HLA class II and I genes had stronger effects in the <7 group, and seven non-HLA regions, one of which functions specifically in beta cells (*GLIS3*), and the other six likely affecting key T cell (*IL2RA, IL10, SIRPG*), thymus (*PTPRK*) and B cell development/functions (*IKZF3, IL10*) or in both immune cells and beta cells (*CTSH*).

Interpretation

In newborn children with the greatest load of certain HLA and non-HLA risk alleles, inherited variants in immune and beta cells, and their inherent disregulated response

51 to environmental stresses such as virus infection, combine to cause a rapid loss of

52 insulin production, thereby driving down the age at which T1D is diagnosed.

Abbreviations

T1D: Type 1 diabetes

HLA: Human leukocyte antigen

FDR: False discovery rate

eQTL: Expression quantitative trait loci

NIDDK: The National Institute of Diabetes and Digestive and Kidney Diseases

NIAID: The National Institute of Allergy and Infectious Diseases

NHGRI: The National Human Genome Research Institute

NICHD: The National Institute of Child Health and Human Development

JDRF: The Juvenile Diabetes Research Foundation

GRID: Genetic resource investigating diabetes

IDDMGEN: Tyypin 1 Diabetekseen Sairastuneita Perheenjäsenineen

T1DGEN: Tyypin 1 Diabetekseen Genetiikka

T1DGC: Type 1 diabetes genetics consortium

IFN: Interferon

Introduction

Type 1 diabetes (T1D) is a multifactorial disease in which the insulin-producing beta cells of pancreatic islets are destroyed or rendered dysfunctional by an autoimmune process that often initiates in the first few months of life, causing a pre-diabetic, non-symptomatic state in approximately 0.4% of children [1]. The actual diagnosis could happen many years after this prodromal phase, the joint environmental and genetic mechanisms of which remain ill defined, with the median age-at-diagnosis being around age 11 years. Even after diagnosis there is still often sufficient endogenous insulin production to lower insulin treatment and reduce the later in life complications of early mortality, cardiovascular, kidney, eye and peripheral neuron disease [2,]. The exceptions to this are the children diagnosed with T1D under the age 10 years in whom there is little insulin production shortly after diagnosis, as measured by circulating C-peptide concentrations [2,3]. This subgroup represents the largest unmet clinical challenge, since they suffer the greatest complications of the disease [3]. Yet any intervention of T1D autoimmunity in these young children must be as safe and precise as possible, modulating the causative molecules, cells, pathways and mechanisms. Hence we need to identify the specific mechanisms underlying early-diagnosed T1D.

Recent evidence suggests that children diagnosed under age 7 years may have a different, more aggressive form of islet inflammation (insulitis), characterised by a B lymphocyte infiltrate coincident with a T cell insulitis ($CD4^+$ and $CD8^+$ T cells), than children aged 13 years and over, who have reduced B cell participation [4]. In cases diagnosed between 7 and 12 years there is a mixture of islet infiltrate phenotypes, some with the “under 7” B cell infiltrate and others with “13 and over” phenotype.

There is already evidence that some genetic variants reduce age-at-diagnosis, which provides insight into the biology of this most beta-cell destructive form of the disease [5–8]. The autoantigen-presenting genes human leukocyte antigen (HLA) class II and class I are the major drivers of younger age-at-diagnosis. Class II molecules are recognised by $CD4^+$ T cells which provide help for $CD8^+$ beta-cell cytotoxic T cells and islet antigen-specific B cells. Class I molecules are expressed on beta cells, upregulated during viral infection or by immune cytokines, rendering them more susceptible to autoreactive $CD8^+$ T cells. More recently, a genome-wide association genetic analysis of age-at-diagnosis of T1D identified a locus on chromosome 6q22.33 that acts almost exclusively in the cases of T1D diagnosed under age 5 years [9], encoding the protein tyrosine phosphatase receptor kappa (*PTPRK*) and thymocyte-expressed molecule involved in selection (*THEMIS*) genes. However, this approach has to meet the stringent genome-wide multiple testing correction criterion ($p < 5 \times 10^{-8}$) and informative, true signals were likely to have been missed. In the present study, we analysed the association of specified known T1D gene regions, thereby reducing the multiple testing burden. In addition, a biological or phenotypic prior could provide greater sensitivity in the search for age-at-diagnosis-associated genes. The stratification of patients into age-at-diagnosis categories according to their pancreatic histology, as opposed to treating age-at-diagnosis as a continuous phenotype provides us with just this opportunity.

Here, we analysed T1D-associated variants according to the proposed pancreatic infiltrate stratification of T1D, namely the age-at-diagnosis groups, the under 7's versus the 13's and over. If T1D has a particular pancreatic immunophenotype then it might be expected that it could have distinct genetic features, characterised by susceptibility genes with larger effects in the under 7's. Moreover, the intermediate

group, age-at-diagnosis 7-13 years, would have risk for these age-at-diagnosis-sensitive genes lying between the under 7's and the 13's and over. Six HLA haplotypes/alleles and seven non-HLA loci fulfil this risk profile informing the biology of the most aggressive form of T1D, revealing a mixture of predisposition in both the beta cell and immune cell compartments.

Methods

Study populations

Our dataset consists of 18,400 controls, 3,111 T1D cases diagnosed at <7 years (the <7 group), 3,759 at ≥7 to <13 years (the 7-13 group) and 1,721 at ≥13 years (the ≥13 group). The majority of individuals are from the UK, with others from central Europe, Asia-Pacific, Finland and the USA (Table 1), and comprises only unrelated individuals, since related individuals were removed (Supplementary methods).

Loci studied

We examined eight HLA class II haplotypes and nine HLA class I classical alleles for their association with T1D diagnosed at each age group, where the haplotypes and classical alleles were a subset of the most protective and susceptible haplotypes identified for T1D to date [10] that we also found to be associated with T1D in our analysis after conditioning for the other associated HLA haplotypes (logistic regression Wald test $p<0.01$). Supplementary Table 1 summarises which haplotypes and classical alleles were examined, how they were defined and whether they were common enough to include in our analysis, defined as at least 5 individuals from each group with the classical allele/haplotype.

We also examined 55 loci outside the HLA, which have previously shown association with T1D (Supplementary Table 2). Each locus contains an 'index' variant, chosen to be the most strongly disease associated from a set of variants in linkage disequilibrium (LD) that constitute a single genetic signal. We have allocated locus names to each of these variants based on a candidate gene(s), but the named genes may not be causal for T1D.

Imputation

Classical HLA alleles as well as non-HLA variants that were excluded due to variant quality control filtering were imputed for analysis (Supplementary methods). Some individuals were genotyped for a subset of their classical HLA alleles [8] and therefore accuracy of imputation was assessed at those classical alleles for a proportion of individuals.

Multinomial logistic regression

In order to examine whether or not there was heterogeneity in effect size for each examined variant between the <7 and ≥13 groups, we fitted two multinomial logistic regressions per locus, one assuming identical effect sizes for the genetic variant in the <7 and ≥13 groups and the other allowing different effect sizes between groups. A comparison of how well these models fit the data allows us to test for heterogeneity in effect size between the two groups. Both models were adjusted for the ten largest principal components derived from the set of ImmunoChip variants passing quality control filters (Supplementary methods).

To test stability of our results at non-HLA loci, we did four sensitivity analyses. Firstly, we sampled without replacement 50% of cases and controls from each of the ancestry groups in our collection. In lieu of a valid replication dataset, to assess the possibility that age-at-diagnosis genetic heterogeneity was due to an unlikely chance distribution of genotypes between age strata, we repeated the heterogeneity test, for the 50% that were sampled and also on the remaining 50%. We performed this procedure 100 times, giving us 200 heterogeneity tests and noted the proportion of times the variant under consideration reached nominally significant heterogeneity ($p<0.05$). Secondly, to exclude the possibility of spurious associations due to population structure in our data, we repeated the analysis but only including individuals from the UK and Northern Ireland and adjusted for the five largest

principal components derived from Immunochip data in these individuals only. Finally, to test sensitivity of our results to age-strata thresholds, we performed the same analysis but instead compared individuals diagnosed at <6 years to the ≥13 group and also individuals diagnosed at <5 years compared to the ≥13 group. We declared a locus differentially-associated if the heterogeneity p-value was associated to a False Discovery Rate (FDR) of <0.1. To explore whether there were more age-at-diagnosis associated variants which we cannot detect in the present analysis due to a lack of statistical power, we examined all loci which did not reach the association threshold (FDR<0.1) and counted how many loci had the largest effect in the <7 group, the intermediate effect in the 7-13 group and the smallest effect in the ≥13 group and compared this to the expected frequency of this ordering using a binomial test (Supplementary methods).

Fine mapping

For each non-HLA locus with strong evidence of heterogeneity between age-at-diagnosis groups, as determined by Bonferroni correction, a more conservative multiple-comparison correction than FDR, we fine mapped a 0.5 Mb region around the index variant to identify a list of potentially causal variants for T1D diagnosed at <7 years. Analysis was limited to individuals from the UK and Northern Ireland, amounting to 2,888 cases diagnosed at <7 years and 11,064 controls, in order to examine a homogeneous population, as fine mapping is sensitive to differences in LD structure between ancestrally divergent groups. We used the GUESSFM software, [11] which carries out a Bayesian variable selection stochastic search to identify the combinations of variants constituting separate genetic susceptibility to T1D. We then examined whether the T1D-associated variants colocalised with expression quantitative trait loci (eQTL) associations in whole blood from a dataset of over

30,000 individuals, gauging which genes the variants are most likely to be regulating and in what direction the effects are on gene transcription and disease risk [12] (eQTL statistics downloaded from http://www.eqtlgen.org/cis-eqtls.html) (Supplementary methods).

The scripts used to analyse these data are available at https://github.com/jinshaw16/AAD_t1d, commit 1727d18c3fe2559ac527681142155b83e8294165.

Funding

This work was funded by the JDRF (9-2011-253, 5-SRA-2015-130-A-N) and Wellcome (091157, 107212) to the Diabetes and Inflammation Laboratory, University of Oxford.

We use data generated by the Wellcome Trust Case Control Consortium (076113). The Northern Irish GRID, IDDMGEN, T1DGEN and Warren cohorts were genotyped using the T1DGC grants from the NIDDK, the NIAID, the NHGRI, the NICHD and the JDRF (U01 DK062418, JDRF 9-2011-530).

Results

Multinomial logistic regression: HLA

We found six HLA haplotypes to be differentially-associated between the <7 and ≥13 group (FDR<0.1). The strongest susceptible class II effect was for the DR3-DQ2/DR4-DQ8 diplotype, whilst the protective DRB1*15:01-DQB1*06:02 and DRB1*07:01-DQB1*03:03 haplotypes showed greater protection from T1D in the <7 group compared to the ≥13 group. Class I alleles A*24:02 and B39*06 showed more susceptibility to T1D in the <7 compared to and ≥13 group (Figure 1). Comparison of imputed classical 4 digit HLA alleles with directly genotyped 4 digit HLA alleles showed concordance of over 91% for each of gene examined (Supplementary Figure 1).

Multinomial logistic regression: non-HLA regions

Outside the HLA, nine regions were differentially-associated between the <7 and ≥13 group (FDR<0.1), near Ikaros family zinc finger 3 (*IKZF3)*, Cathepsin H (*CTSH*), GLIS family zinc finger 3 (*GLIS3*), Chymotrypsinogen B1 (*CTRB1)*, the third index variant at interleukin 2 receptor alpha (*IL2RA*), interleukin 10 (*IL10*), Calmodulin-Regulated Spectrin-Associated Protein 2 (*CAMSAP2*), Signal Regulatory Protein Gamma (*SIRPG*) and *PTPRK* (Figure 2). Three of these (*IKZF3*, *CTSH* and *GLIS3*) survived Bonferroni correction (p<0.05/55=0.00091). At each locus associated with FDR<0.1, the 7-13 group had a larger effect size than the ≥13 group and smaller than the <7 group. Given the ≥13 group comprises just 1,721 individuals, it is probable that with increased sample size and hence statistical power, other T1D risk loci with sizeable estimated effect size differences between groups might reach statistical significance with regards to heterogeneity (Supplementary Figure 2). Of the 46 variants not satisfying an FDR<0.1, 20 have the strongest signal in <7s, weakest in

≥13s and intermediate in 7s-13s, compared to 8 occurrences in that order expected by chance ($p=4.27\times10^{-6}$, binomial test), suggesting the presence of substantial additional signal in variants that we are not able to declare show evidence individually.

Stability of non-HLA results

After sampling half of the cases and controls 100 times, we found that nominal significance ($p<0.05$) was observed for the heterogeneity in effect size test between age-at-diagnosis groups >50% of the time for the *IKZF3*, *CTSH*, *GLIS3*, *CTRB1* and *IL2RA* (3rd index variant) loci (Supplementary Figure 3), and >44% of the time at the other FDR heterogeneous loci (*IL10*, *CAMSAP2*, *SIRPG* and *PTPRK/THEMIS*).

In the UK-specific sensitivity analysis, six of the nine FDR heterogeneous loci from the primary analysis were heterogeneous between the <7 and ≥13 group (FDR<0.1) in this ancestry-homogeneous population, two of the loci showed no heterogeneity in effect size (*CTRB1* p=0.310 and *CAMSAP2* p=0.578) and were thus removed from our set of differentially-associated regions and the remaining locus, *IL10*, had a p-value of 0.06, which we considered differentially associated between the <7 and ≥13 groups, given the decrease in statistical power in this sensitivity analysis (Supplementary Figure 4).

When changing the threshold for the early-diagnosed group to <6 and <5, all seven associated loci from the primary analysis and UK-specific analysis were heterogeneous (FDR<0.1) (Supplementary Figures 5 and 6).

Minor allele frequency plots by age-at-diagnosis for the seven differentially associated loci that remained heterogeneous between the <7 and ≥13 group in all analyses are shown in Supplementary Figures 7-13, whilst Supplementary Tables 3 and 4 summarise the most likely causal genes at these loci.

Fine mapping

We fine mapped the three loci (*IKZF3, CTSH* and *GLIS3*) that reached Bonferroni-corrected heterogeneity between age-at-diagnosis groups. The posterior probability of there being one causal variant was >0.63 at each locus. All variants contained within a group that has a group posterior probability of causality of >0.9 are listed in Supplementary Tables 5-7, though our stringent post-imputation variant quality control filtering means variants in high LD with the listed variants could also be causal for T1D but were removed from the current analysis due to low quality imputation at that variant.

The *IKZF3* locus results prioritise an LD block containing 34 variants, all of which could be causal, which also effects expression of at least three genes ($p<5 \times 10^{-150}$), where the minor allele at the most likely causal variants decrease T1D risk and *IKZF3* expression and also increase expression of *GSDMB* and *ORMDL3*. Colocalisation analyses support the hypothesis that the disease causal variant and the eQTL causal variant were the same for all three genes (posterior probability of colocalisation for T1D and eQTL with *IKZF3*=0.973, *GSDMB*=0.841 and *ORMDL3*=0.846).

The *CTSH* locus showed evidence of colocalisation with the *CTSH* whole blood eQTL (posterior probability of colocalisation=0.655); the susceptibility allele for T1D is associated with more expression of *CTSH* (Figure 3).

There was no evidence of colocalisation between disease risk and *GLIS3* whole blood eQTL (posterior probability of colocalisation=0.036), suggesting the variant might be acting elsewhere to alter T1D risk.

Discussion

The stratification of patients by age-at-diagnosis according to islet phenotypes has provided a rich source of genes, molecules and pathways with greater effects in children diagnosed with T1D under age 7 years. We expected to see strong differential associations with the HLA class II haplotypes, in particular the strongest single susceptibility determinant in the genome, the heterozygous diplotype DR3-DQ2/DR4-DQ8. Previously with smaller sample sizes and without dichotomising patients into biologically-defined discrete age categories, HLA class I alleles, A*24:02 and B*39:06 have been shown to be associated with younger age-at-diagnosis [5–8]. Here, we show for the first time that the protective HLA class II haplotypes DRB1*15:01-DQB1*06:02 and DRB1*07:01-DQB1*03:03 are less prevalent amongst individuals diagnosed at <7 years compared with controls and those diagnosed at ≥13 years. Therefore, the earliest and most aggressive phenotypic subtype of T1D results primarily from carriage of high risk alleles and haplotypes of the HLA class II and I genes, which probably act at four levels: (i) altering the T cell receptor repertoire in favour of anti-islet antigen reactivity, for example preproinsulin, and/or reducing the protective repertoire of T regulatory cells; (ii) providing a strong autoantigen presentation environment in the islets and pancreatic draining lymph nodes enabling the infiltration and cytolytic activity of $CD8^+$ T cells but also by disrupting B cell anergy [13] permitting binding and presentation of autoantigen to provide potent help to T cells in a self-reinforcing spiral of autoreactivity; (iii) affecting the immune response to the viral infections that are involved in the disease; (iv) affecting how the gut microbiome develops in early life, a system that is known to affect T1D susceptibility [14].

In addition to the HLA heterogeneity, we obtained robust evidence of differences in effect size between the age-at-diagnosis groups at seven non-HLA loci. Of these loci, one plausible candidate gene, *GLIS3*, most likely perturbs disease risk in the islet beta cells, given the expression levels in the pancreas, lack of expression in immune cells, colocalisation with type 2 diabetes risk variants [15] and lack of association with other autoimmune diseases (https://genetics.opentargets.org). This finding supports a mechanism of beta-cell fragility, for example, susceptibility to apoptosis [16], in which increased risk of disease is encoded in the beta cell, not only in the immune system. The *GLIS3* effect can be mimicked in a mouse model of non-immune diabetes by a high fat diet, linking obesity as a risk factor in T1D and type 2 diabetes [16,17]. Two of the loci, *CTSH* and *IKZF3*, could act in the islets or elsewhere, whilst all of the other candidate causal genes (*IL2RA*, *IL10*, *SIRPG*, *PTPRK/THEMIS*, as well as *IKZF3/ORMDL3/GSDMB* and *CTSH*) have known functions in T and/or B cell biology (Supplementary Table 4). This implies that in addition to HLA-susceptibility, risk of T1D in the very young is also impacted by particular malfunctions in the infiltrating T and B cells, leading to increased risk of autoreactivity, resulting in a perfect storm of immune infiltration, antigen recognition and a rapid destruction of beta cells.

Of the three non-HLA risk regions with the strongest evidence of heterogeneity between age-at-diagnosis groups, we focus on the *IKZF3* and *CTSH* loci, which colocalise with whole blood eQTLs. The region containing *IKZF3* has a complex structure with a large LD block, which is associated with multiple diseases, including asthma and paediatric asthma [18, 19] (Supplementary Table 4). However, the direction of effect of the risk variant is opposite in asthma to all associated autoimmune diseases, including T1D, where the C allele at a variant within the haplotype,

rs921649 (C>T), increases susceptibility to autoimmunity, whereas the C allele is protective for asthma [18]. Whole blood eQTL data shows the expression of 13 protein-coding genes is modulated by variants in the disease-associated haplotype, with *IKZF3*, *ORMDL3* and *GSDMB* the most affected [12]. All three genes are expressed in lymphocytes and are up- (*IKZF3*) or down-regulated (*ORMDL3*, *GSDMB*) (https://dice-database.org/) following activation, with good biological candidacy for altering disease risk. *IKZF3* is a transcriptional repressor with a key role in B-cell activation and differentiation [20] and T cell differentiation [21]. *ORMDL3* is a central regulator of sphingolipid biosynthesis [22] and has also been proposed to negatively regulate store-operated calcium, lymphocyte activation and cytokine production [18,23], while *GSDMB* can act as a pyroptotic protein [24]. Therefore one or more of these genes may be causal for T1D risk. Pertinent to the increased frequency of B-cell infiltration in the islets of the <7 group, there is evidence that carriers of the T1D risk allele have decreased anergic high affinity insulin-binding B cells in circulating blood, implying some of this population may have relocated to the pancreas [13]. This loss of anergic circulating B cell frequencies is also associated with the most predisposing age-at-diagnosis diplotype HLA-DRB1*03:01-DQB1*02:01/DRB1*04:01-DQB1*03:02 compared to donors with the protective HLA class II haplotypes [13].

The candidate T1D risk variants at the *CTSH* locus, for example the C allele at rs2289702 (C>T), are associated with increased expression of CTSH RNA in multiple cell types and tissues (Supplementary Table 4). The locus has previously been implicated in T1D aetiology by altering sensitivity of beta cells to apoptosis [25], where rs3825932 (C>T) was investigated, which is in low LD (r^2=0.26) with the disease-associated variant reported here and the T1D risk allele counter-intuitively resulted in protection from beta-cell apoptosis. Thus, beta-cell apoptosis may not be the primary

356 mechanism underlying disease aetiology in this region. *CTSH* functions as an
357 endopeptidase and can cleave the N-terminus of the Toll-like receptor 3 (TLR3)
358 protein, increasing its functionality [26]. Given TLR3 is expressed in islets [27], it is
359 possible that the increase in CTSH expression associated with the T1D susceptibility
360 allele (the C allele of rs2289702) results in increased TLR3 N-terminus cleavage,
361 heightened responses to viral infections and increased release of type 1 interferon
362 (IFN). This may increase baseline risk of T1D and specifically the risk of early-
363 diagnosed T1D in individuals carrying this allele, since viral infections are more
364 frequent in childhood. There is mounting evidence that enteroviral infections
365 predispose to T1D, a type 1 IFN transcriptional signature precedes anti-islet
366 autoantibody appearance in children [28], and another receptor for viral RNA, MDA5
367 encoded by *IFIH1* is a proven T1D susceptibility gene with its higher IFN-inducing
368 activity increasing risk of the disease [29]. Exposure of beta cells to type 1 IFN greatly
369 increases their HLA class I expression and susceptibility to $CD8^+$ cytotoxic killing,
370 and heightened class I expression on beta cells is a hallmark phenotype of the T1D
371 pancreas [30].
372 Our genetic results imply a dynamic, fully integrated pathogenic collaboration
373 between the immune system, the beta cells and viral infection in the initiation and
374 rapid development of extreme insulin-deficiency starting in the first few weeks and
375 months of life in those that carry the heaviest load of age-at-diagnosis alleles.
376 Combinations of modulators of these pathways could be an effective way of
377 preventing the cessation of endogenous insulin-production.

Tables and Figures

Table 1: Characteristics of individuals included in the analysis.

	Controls	**<7**	**7-13**	**>13**
N	18400	3111	3759	1721
Mean age-at-diagnosis	-	3.67	9.54	18.31
Sex: Female	9712 (52.8%)	1495 (48.8%)	1917 (52%)	638 (43%)
Asia-Pacific	858 (4.7%)	22 (0.7%)	24 (0.6%)	34 (2%)
Central Europe	1680 (9.1%)	50 (1.6%)	62 (1.6%)	172 (10%)
Finland	2819 (15.3%)	100 (3.2%)	154 (4.1%)	438 (25.5%)
Northern Ireland	478 (2.6%)	222 (7.1%)	248 (6.6%)	35 (2%)
UK	10586 (57.5%)	2666 (85.7%)	3209 (85.4%)	922 (53.6%)
USA	1979 (10.8%)	51 (1.6%)	62 (1.6%)	120 (7%)

Figure 1: Classical HLA alleles association with type 1 diabetes diagnosed at <7 years old (red circle; mean log-odds ratio age-at-diagnosis +/- 95%CI), 7-13 years old (green circle; mean log-odds ratio age-at-diagnosis 7-13+/- 95%CI) and ≥13 years old (blue circle mean log-odds ratio age-at-diagnosis >13+/- 95%CI), from a multinomial logistic regression. Left panel shows the log-odds ratios with a dashed red line showing a log-odds ratio of 0. The right panel shows the association statistics from a likelihood ratio test comparing a multinomial logistic regression constraining the log-odds ratios from the <7 and ≥13 groups to be equal compared to an unconstrained model. Red dotted line shows nominal significance in heterogeneity ($p<0.05$), red dashed line show Bonferroni-corrected significance in heterogeneity.

Figure 2: Non-HLA type 1 diabetes associated variants, showing log-odds ratios for those diagnosed at <7 years old (red circle; mean log-odds ratio age-at-diagnosis +/- 95%CI), 7-13 years old (green circle; mean log-odds ratio age-at-diagnosis 7-13+/- 95%CI) and ≥13 years old (blue circle mean log-odds ratio age-at-diagnosis ≥13+/- 95%CI), from a multinomial logistic regression. Left panel shows the log-odds ratios with a dashed red line showing a log-odds ratio of 0. The right panel shows the association statistics from a likelihood ratio test comparing a multinomial logistic regression constraining the log-odds ratios from the <7 and ≥13 groups to be equal compared to an unconstrained model. Showing only loci with a false discovery rate of less than 0.1. Red dotted line shows threshold for false discovery rate of <0.1, red dashed line shows threshold for Bonferroni-corrected heterogeneity.

Figure 3: Results from fine mapping the *IKZF3* and *CTSH* loci regions. Analysis includes individuals from the UK and Northern Ireland only and only controls and

406 cases diagnosed at <7 years. Top panel: Gene positions (genome build 37), with
407 arrows indicating direction of transcription. Second panel, univariable early-
408 diagnosed (<7) type 1 diabetes log-odds ratios and 95% confidence intervals for each
409 of the most likely causally associated variants as prioritised by GUESSFM. Third
410 panel: $\log_e$(absolute eQTL z score) if z score>0 and -$\log_e$(absolute eQTL z score) if z
411 score<0, so direction of effect can be compared, from publically available whole
412 blood eQTL dataset of over 30,000 individuals, including only eQTLs with a p value
413 of $<5\times10^{-150}$ (*IKZF3* locus) and $<5\times10^{-50}$ (*CTSH* locus). The symbols are coloured red
414 if contained in the set of most likely causal variants, as produced by GUESSFM and
415 the shape corresponds to the gene that the variant is effecting transcription of.
416
417 **Supplementary Figure 1:** Concordance of HiBag imputed versus directly genotyped
418 classical HLA alleles. Concordance is defined as identical 4 digit HLA classical allele
419 at both chromosomes.
420
421 **Supplementary Figure 2:** All 55 non-HLA type 1 diabetes associated variants,
422 showing log-odds ratios for those diagnosed at <7 years old (red circle; mean log-
423 odds ratio age-at-diagnosis +/- 95%CI), 7-13 years old (green circle; mean log-odds
424 ratio age-at-diagnosis 7-13+/- 95%CI) and ≥13 years old (blue circle mean log-odds
425 ratio age-at-diagnosis ≥13+/- 95%CI), from a multinomial logistic regression. Left
426 panel shows the log-odds ratios with a dashed red line showing a log-odds ratio of 0.
427 The right panel shows the association statistics from a likelihood ratio test comparing
428 a multinomial logistic regression constraining the log-odds ratios from the <7 and ≥13
429 groups to be equal compared to an unconstrained model. Red dotted line shows
430 threshold for nominally significant heterogeneity between groups ($p<0.05$), red solid

line shows threshold for false discovery rate of <0.1, red dashed line shows threshold for Bonferroni-corrected significant heterogeneity.

Supplementary Figure 3: Proportion of times each locus had a nominally significant heterogeneity test ($p<0.05$) when sampling 50% of the entire collection 100 times, generating 200 analysis datasets.

Supplementary Figure 4: All 55 non-HLA type 1 diabetes associated variants examined using only individuals from the UK or Northern Ireland, showing log-odds ratios for those diagnosed at <7 years old (red circle; mean log-odds ratio age-at-diagnosis +/- 95%CI), 7-13 years old (green circle; mean log-odds ratio age-at-diagnosis 7-13+/- 95%CI) and ≥13 years old (blue circle mean log-odds ratio age-at-diagnosis >13+/- 95%CI), from a multinomial logistic regression. Left panel shows the log-odds ratios with a dashed red line showing a log-odds ratio of 0. The right panel shows the association statistics from a likelihood ratio test comparing a multinomial logistic regression constraining the log-odds ratios from the <7 and ≥13 groups to be equal compared to an unconstrained model. Red dotted line shows threshold for nominally significant heterogeneity between groups ($p<0.05$), red solid line shows threshold for false discovery rate of <0.1, red dashed line shows threshold for Bonferroni-corrected significant heterogeneity.

Supplementary Figure 5: All 55 non-HLA type 1 diabetes associated variants, showing log-odds ratios for those diagnosed at <6 years old (red circle; mean log-odds ratio age-at-diagnosis +/- 95%CI), 6-13 years old (green circle; mean log-odds ratio age-at-diagnosis 6-13+/- 95%CI) and ≥13 years old (blue circle mean log-odds

ratio age-at-diagnosis ≥13+/- 95%CI) from a multinomial logistic regression. Left panel shows the log-odds ratios with a dashed red line showing a log-odds ratio of 0. The right panel shows the association statistics from a likelihood ratio test comparing a multinomial logistic regression constraining the log-odds ratios from the <6 and ≥13 groups to be equal compared to an unconstrained model. Red dotted line shows threshold for nominally significant heterogeneity between groups ($p<0.05$), red solid line shows threshold for false discovery rate of <0.1, red dashed line shows threshold for Bonferroni-corrected significant heterogeneity.

Supplementary Figure 6: All 55 non-HLA type 1 diabetes associated variants, showing log-odds ratios for those diagnosed at <5 years old (red circle; mean log-odds ratio age-at-diagnosis +/- 95%CI), 5-13 years old (green circle; mean log-odds ratio age-at-diagnosis 5-13+/- 95%CI) and ≥13 years old (blue circle mean log-odds ratio age-at-diagnosis ≥13+/- 95%CI) from a multinomial logistic regression. Left panel shows the log-odds ratios with a dashed red line showing a log-odds ratio of 0. The right panel shows the association statistics from a likelihood ratio test comparing a multinomial logistic regression constraining the log-odds ratios from the <7 and ≥13 groups to be equal compared to an unconstrained model. Red dotted line shows threshold for nominally significant heterogeneity between groups ($p<0.05$), red solid line shows threshold for false discovery rate of <0.1, red dashed line shows threshold for Bonferroni-corrected significant heterogeneity.

Supplementary Figure 7: Minor allele frequency of the index variant near the *IKZF3* gene for controls and individuals diagnosed at various ages.

Supplementary Figure 8: Minor allele frequency of the index variant near the *CTSH* gene for controls and individuals diagnosed at various ages.

Supplementary Figure 9: Minor allele frequency of the index variant near the *GLIS3* gene for controls and individuals diagnosed at various ages.

Supplementary Figure 10: Minor allele frequency of the index variant near the *IL2RA* gene (third index variant) for controls and individuals diagnosed at various ages.

Supplementary Figure 11: Minor allele frequency of the index variant near the *IL10* gene for controls and individuals diagnosed at various ages.

Supplementary Figure 12: Minor allele frequency of the index variant near the *SIRPG* gene for controls and individuals diagnosed at various ages.

Supplementary Figure 13: Minor allele frequency of the index variant near the *PTPRK/THEMIS* genes for controls and individuals diagnosed at various ages.

Supplementary Table 1: Classical HLA alleles/haplotypes examined in analysis.

Supplementary Table 2: Non-HLA variants examined in analysis.

Supplementary Table 3: Non-HLA region variants with evidence of heterogeneity in effect size between the <7 and ≥13 groups: Promoter Capture Hi-C (PCHi-C) candidate genes.

Supplementary Table 4: Details of non-HLA variants with evidence of heterogeneity in effect size between the <7 and ≥13 groups.

Supplementary Table 5: Most likely variants causally associated with T1D at the *IKZF3* locus from GUESSFM fine mapping analysis.

Supplementary Table 6: Most likely variants causally associated with T1D at the *CTSH* locus from GUESSFM fine mapping analysis.

Supplementary Table 7: Most likely variants causally associated with T1D at the *GLIS3* locus from GUESSFM fine mapping analysis.

519 1. Raab, J. *et al.* Capillary blood islet autoantibody screening for identifying pre-
520 type 1 diabetes in the general population: design and initial results of the Fr1da
521 study. *BMJ Open* **6,** e011144 (2016).
522 2. Kuhtreiber, W. M. *et al.* Low levels of C-peptide have clinical significance for
523 established Type 1 diabetes. *Diabet. Med.* **32,** 1346–1353 (2015).
524 3. Rawshani, A. *et al.* Excess mortality and cardiovascular disease in young
525 adults with type 1 diabetes in relation to age at onset : a nationwide , register-
526 based cohort study. *Lancet* **392,** 477–486 (2018).
527 4. Leete, P. *et al.* Differential insulitic profiles determine the extent of β-cell
528 destruction and the age at onset of type 1 diabetes. *Diabetes* **65,** 1362–1369
529 (2016).
530 5. Howson, J. M. M. *et al.* Evidence of gene-gene interaction and age-at-
531 diagnosis effects in type 1 diabetes. *Diabetes* **61,** 3012–3017 (2012).
532 6. Valdes, A. M. *et al.* Use of class I and class II HLA loci for predicting age at
533 onset of type 1 diabetes in multiple populations. *Diabetologia* **55,** 2394–2401
534 (2012).
535 7. Nejentsev, S. *et al.* Localization of type 1 diabetes susceptibility to the MHC
536 class I genes HLA-B and HLA-A. *Nature* **450,** 887–892 (2007).
537 8. Howson, J. M. M., Walker, N. M., Clayton, D. & Todd, J. A. Confirmation of
538 HLA class II independent type 1 diabetes associations in the major
539 histocompatibility complex including HLA-B and HLA-A. *Diabetes, Obes.*
540 *Metab.* **11,** 31–45 (2009).
541 9. Inshaw, J. R. J., Walker, N. M., Wallace, C., Bottolo, L. & Todd, J. A. The
542 chromosome 6q22.33 region is associated with age at diagnosis of type 1
543 diabetes and disease risk in those diagnosed under 5 years of age. *Diabetologia*

61, 147–157 (2018).

10. Noble, J. A. & Valdes, A. M. Genetics of the HLA Region in the Prediction of Type 1 Diabetes. *Curr. Diab. Rep.* **11,** 533–542 (2011).

11. Wallace, C. *et al.* Dissection of a Complex Disease Susceptibility Region Using a Bayesian Stochastic Search Approach to Fine Mapping. *PLOS Genet.* **11,** e1005272 (2015).

12. Võsa, U. *et al.* Unraveling the polygenic architecture of complex traits using blood eQTL meta- analysis. *bioRxiv* 1–57 (2018).

13. Smith, M. J. *et al.* Loss of B-Cell Anergy in Type 1 Diabetes Is Associated With High-Risk HLA and Non-HLA Disease Susceptibility Alleles. *Diabetes* **67,** 697–703 (2018).

14. Paun, A. *et al.* Association of HLA-dependent islet autoimmunity with systemic antibody responses to intestinal commensal bacteria in children. *Sci. Immunol.* **4,** eaau8125 (2019).

15. Aylward, A., Chiou, J., Okino, M., Kadakia, N. & Gaulton, K. J. Shared genetic contribution to type 1 and type 2 diabetes risk. *BioRxiv* (2018).

16. Dooley, J. *et al.* Genetic predisposition for beta cell fragility underlies type 1 and type 2 diabetes. *Nat. Genet.* **48,** 519–527 (2016).

17. Nogueira, T. C. *et al.* GLIS3, a Susceptibility Gene for Type 1 and Type 2 Diabetes , Modulates Pancreatic Beta Cell Apoptosis via Regulation of a Splice Variant of the BH3-Only Protein Bim. *PLoS Genet.* **9,** e1003532 (2013).

18. Schmiedel, B. J. *et al.* 17q21 asthma-risk variants switch CTCF binding and regulate IL-2 production by T cells. *Nat. Commun.* **7,** (2016).

19. Bouzigon, E. *et al.* Effect of 17q21 Variants and Smoking Exposure in Early-Onset Asthma. *N. Engl. J. Med.* **359,** 1985–1994 (2019).

20. Wang, J. *et al.* Aiolos Regulates B Cell Activation and Maturation to Effector State. *Immunity* **9,** 543–553 (1998).

21. Quintana, F. J. *et al.* Aiolos promotes TH17 differentiation by directly silencing Il2 expression. *Nat. Immunol.* **13,** 770–777 (2012).

22. Davis, D., Kannan, M. & Wattenberg, B. Orm / ORMDL proteins: Gate guardians and master regulators. *Adv. Biol. Regul.* **70,** 3–18 (2018).

23. Carreras-Sureda, A. *et al.* ORMDL3 modulates store-operated calcium entry and lymphocyte activation. *Hum. Mol. Genet.* **22,** 519–530 (2013).

24. Panganiban, R. A. *et al.* A functional splice variant associated with decreased asthma risk abolishes the ability of gasdermin B to induce epithelial cell pyroptosis. *J. Allergy Clin. Immunol.* **142,** 1469–1478 (2018).

25. Fløyel, T., Brorsson, C., Nielsen, L. B., Miani, M. & Bang-berthelsen, C. H. CTSH regulates β -cell function and disease progression in newly diagnosed type 1 diabetes patients. *PNAS* **111,** 10305–10310 (2014).

26. Qi, R., Singh, D. & Kao, C. C. Proteolytic Processing Regulates Toll-like Receptor 3 Stability and Endosomal Localization. *J. Biol. Chem.* **287,** 32617–32629 (2012).

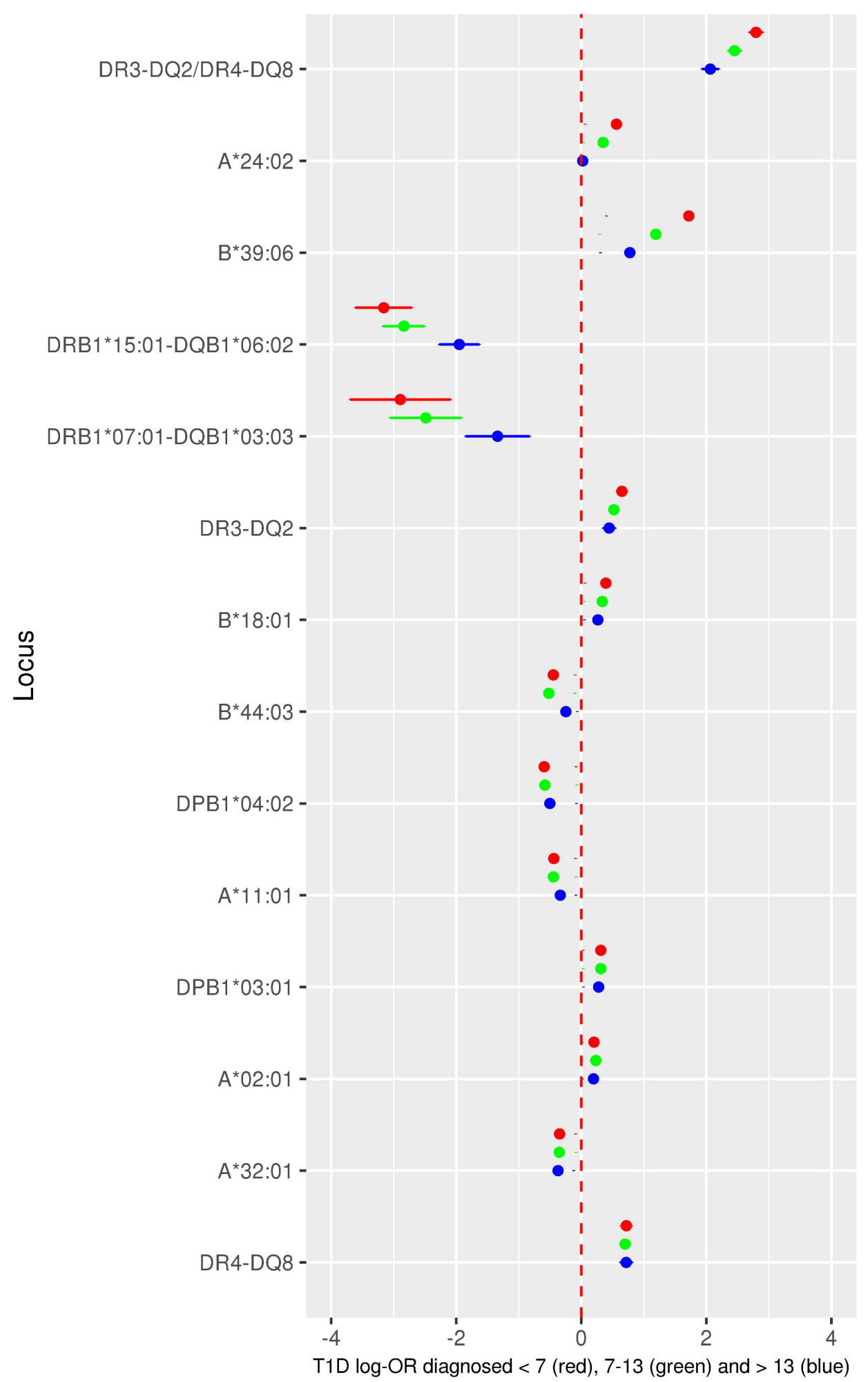
27. Rasschaert, J. *et al.* Toll-like Receptor 3 and STAT-1 Contribute to Double-stranded RNA+ Interferon-gamma-induced Apoptosis in Primary Pancreatic beta-Cells. *J. Biol. Chem.* **280,** 33984–33991 (2005).

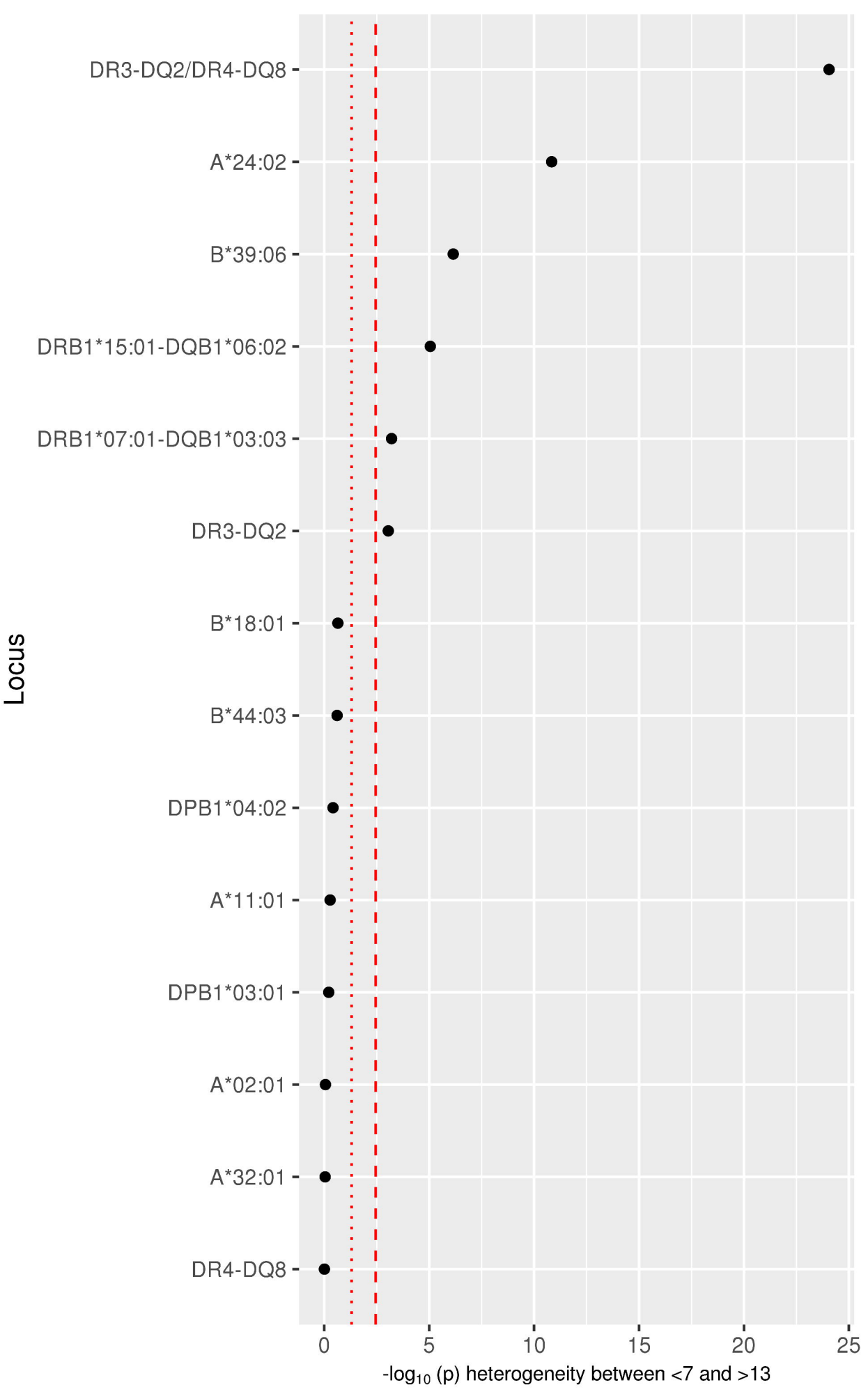
28. Ferreira, R. C. *et al.* A Type 1 Interferon Transcriptional Signature Precedes Autoimmunity in Children Genetically at Risk for Type 1 Diabetes. *Diabetes* **63,** 2538–2550 (2014).

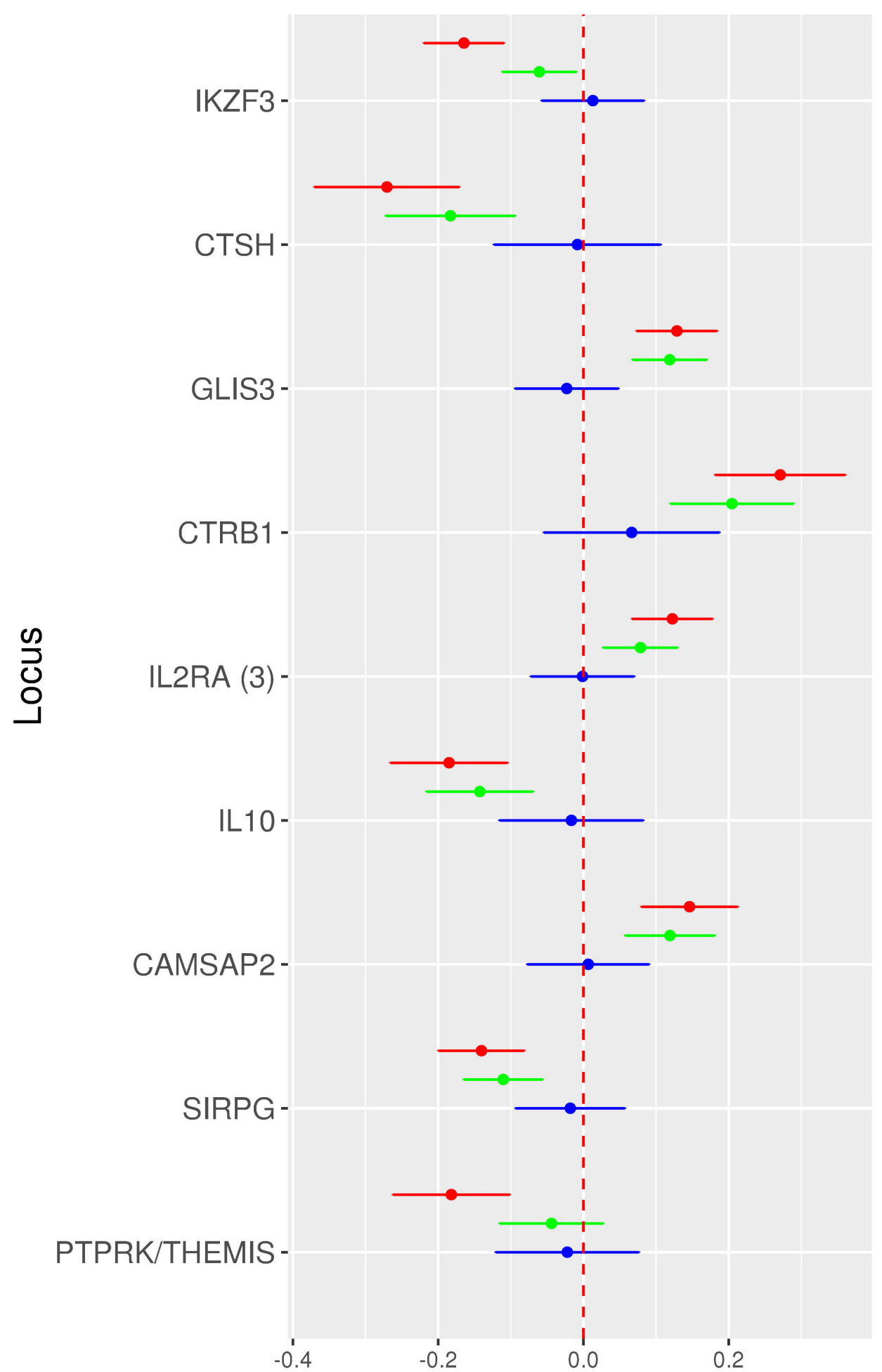
29. Helicase, A. I. *et al.* An Interferon-Induced Helicase (IFIH1) Gene Polymorphism Associates With Different Rates of Progression From

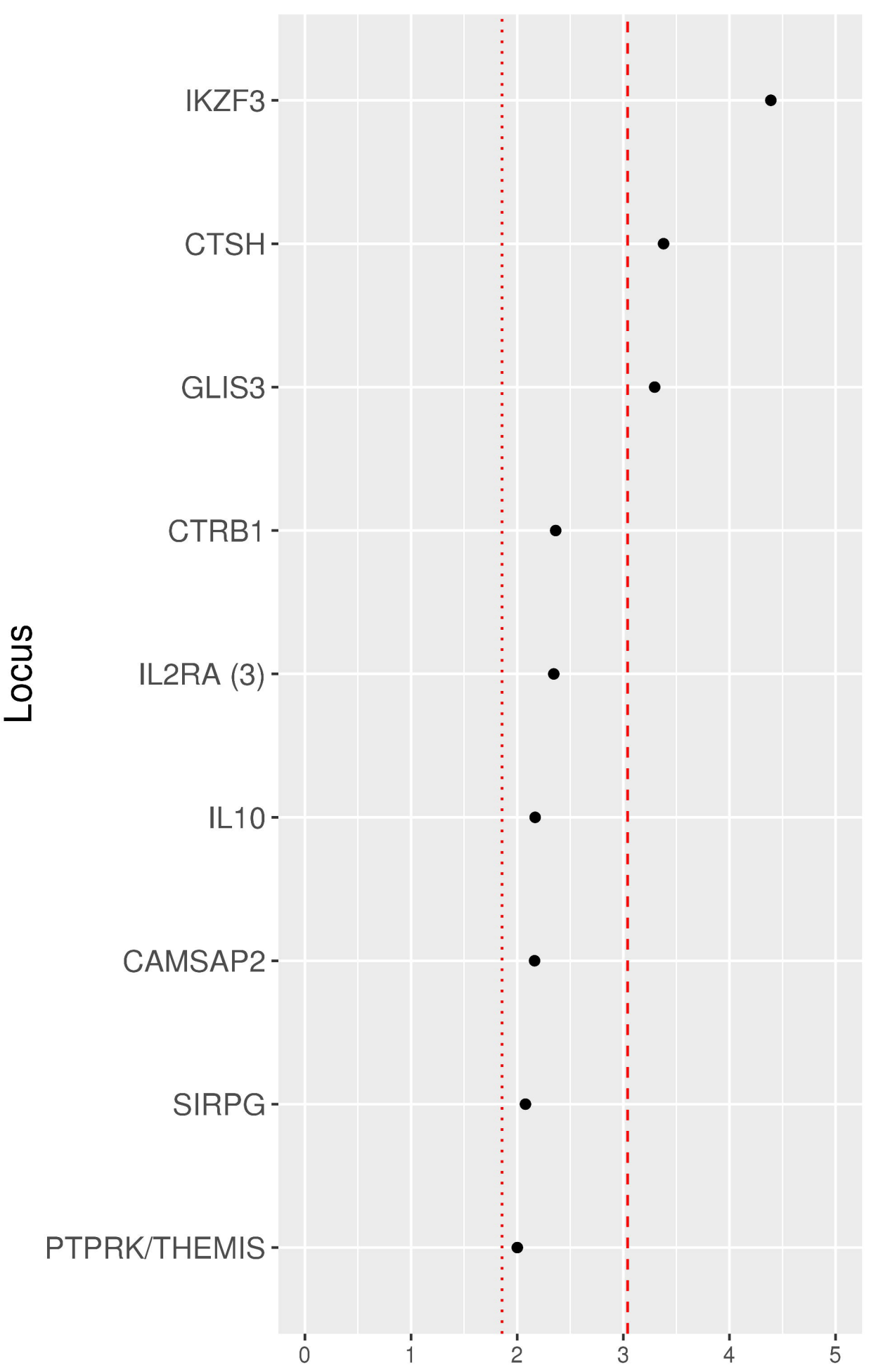
Autoimmunity to Type 1 Diabetes. *Diabetes* **60,** 685–690 (2011).

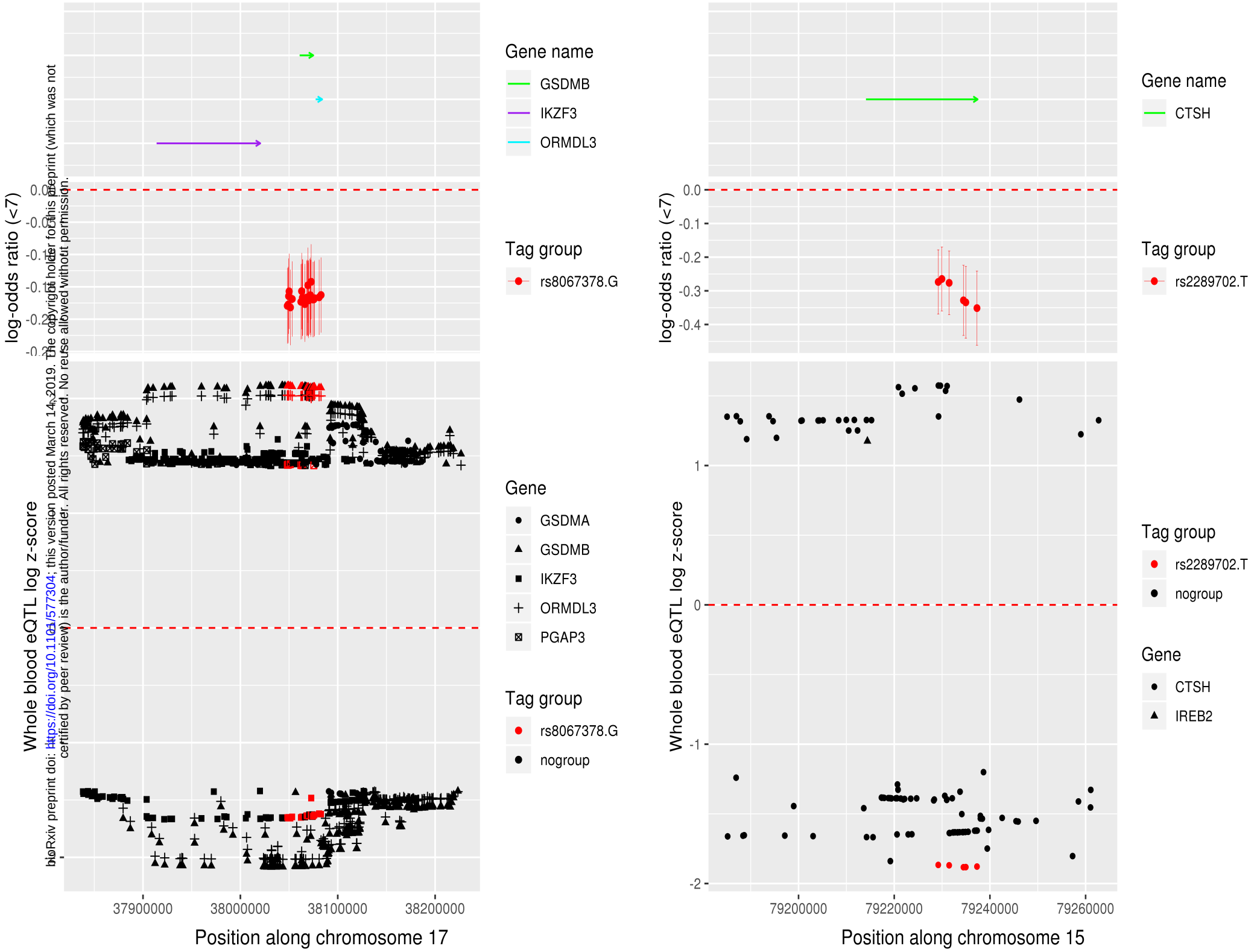
30. Richardson, S. J. *et al.* Islet cell hyperexpression of HLA class I antigens: a defining feature in type 1 diabetes. *Diabetologia* **59,** 2448–2458 (2016).

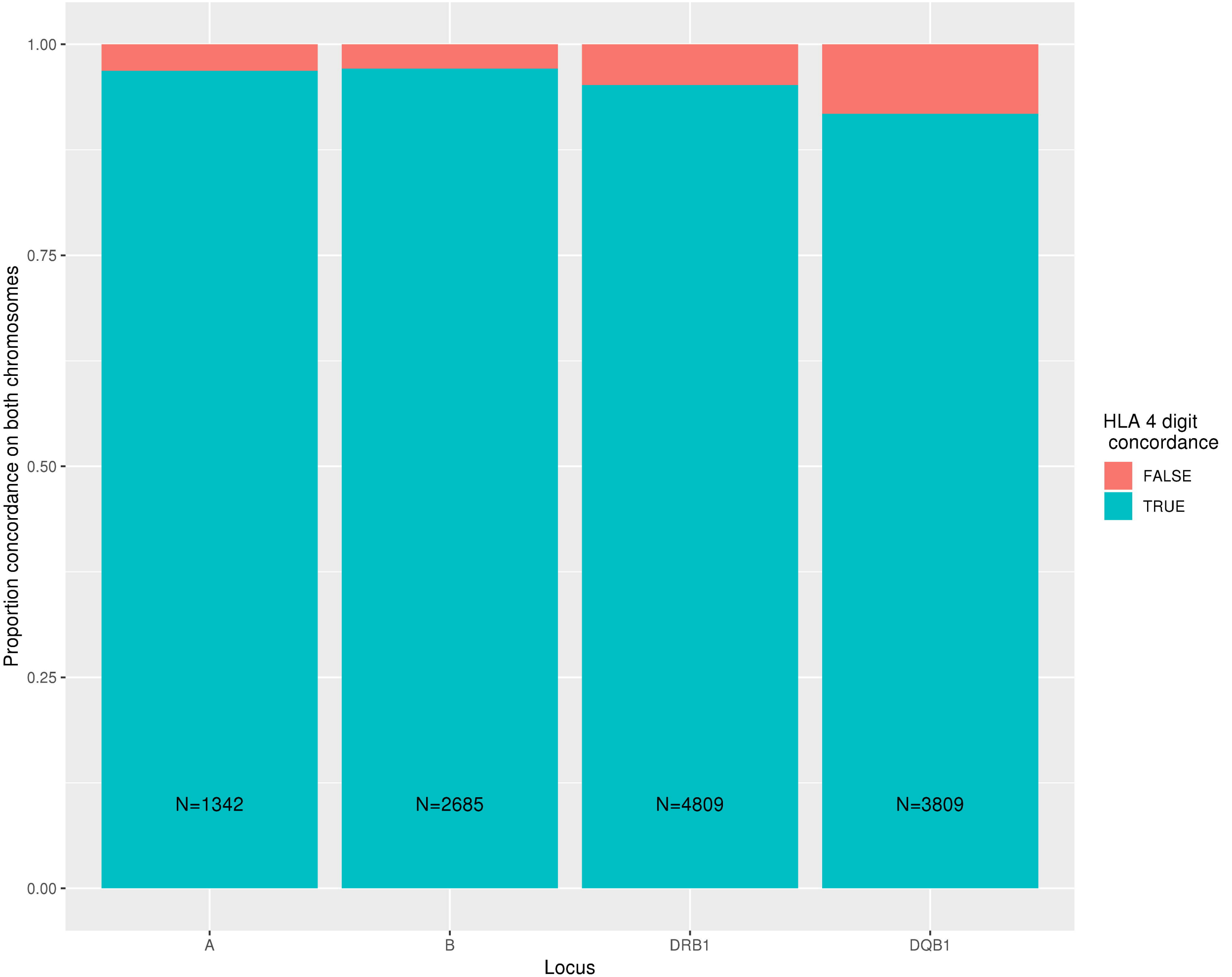

DR3-DQ2/DR4-DQ8
A*24:02
B*39:06
DRB1*15:01-DQB1*06:02
DRB1*07:01-DQB1*03:03
DR3-DQ2
B*18:01
B*44:03
DPB1*04:02
A*11:01
DPB1*03:01
A*02:01
A*32:01
DR4-DQ8
Locus
-4
-2
0
2
4
T1D log-OR diagnosed < 7 (red), 7-13 (green) and > 13 (blue)

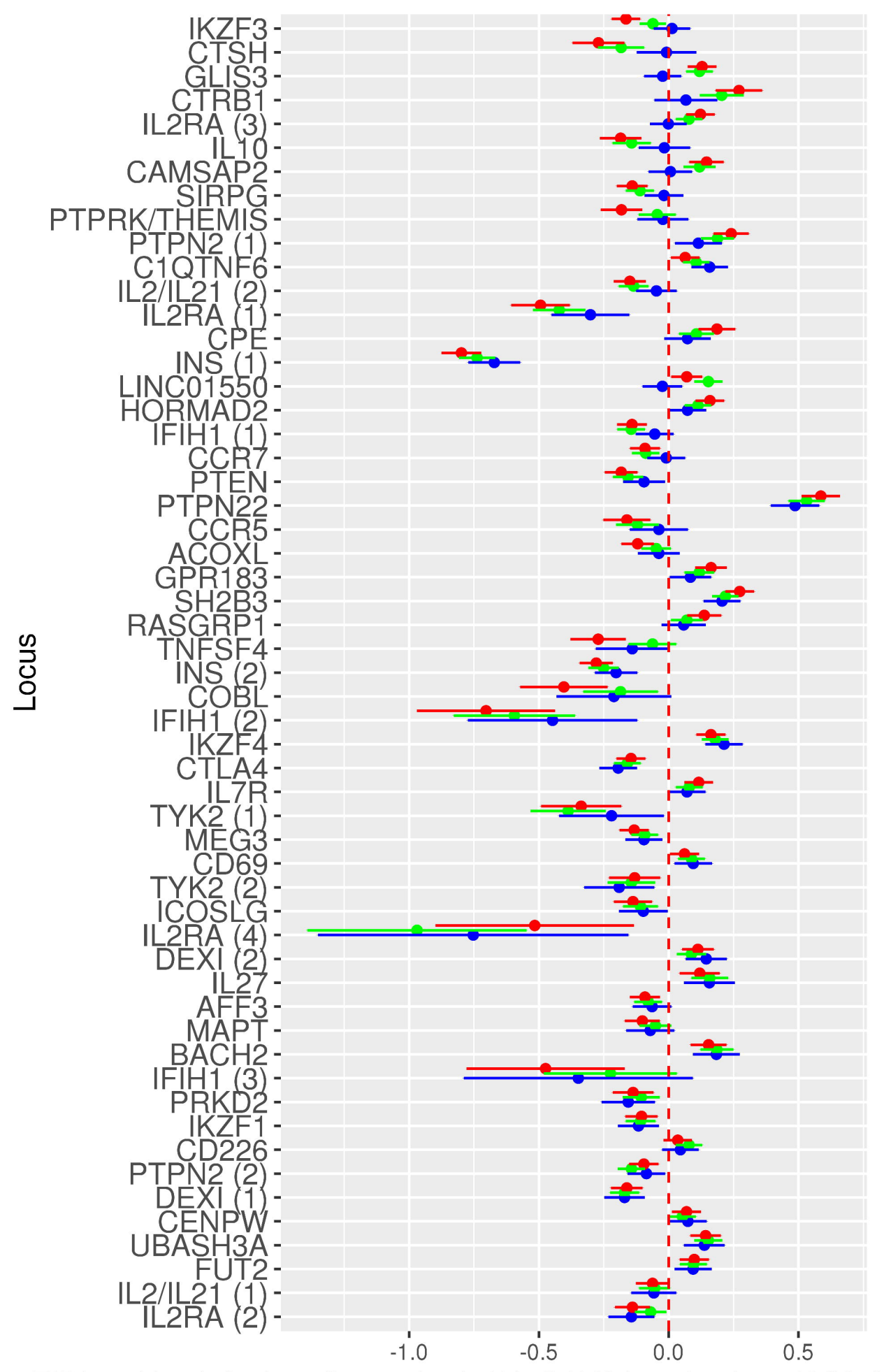

DR3-DQ2/DR4-DQ8
A*24:02
B*39:06
DRB1*15:01-DQB1*06:02
DRB1*07:01-DQB1*03:03
DR3-DQ2
B*18:01
B*44:03
DPB1*04:02
A*11:01
DPB1*03:01
A*02:01
A*32:01
DR4-DQ8
Locus
0
5
10
15
20
25
-log10 (p) heterogeneity between <7 and >13

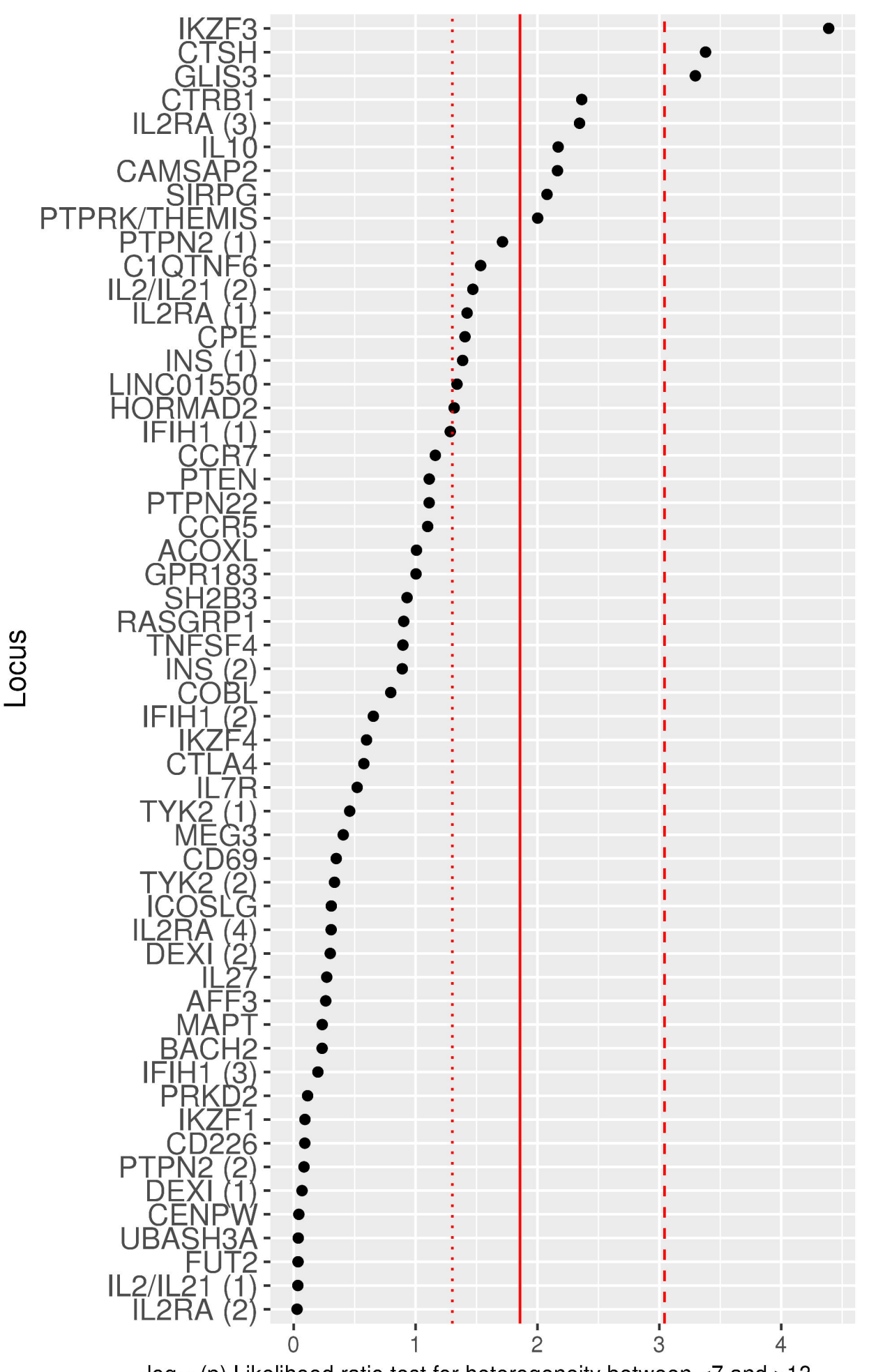

Locus
IKZF3
CTSH
GLIS3
CTRB1
IL2RA (3)
IL10
CAMSAP2
SIRPG
PTPRK/THEMIS
-0.4
-0.2
0.0
0.2
T1D log-odds ratio for those diagnosed under 7 (red), 7-13 (green) and over 13 (blue)

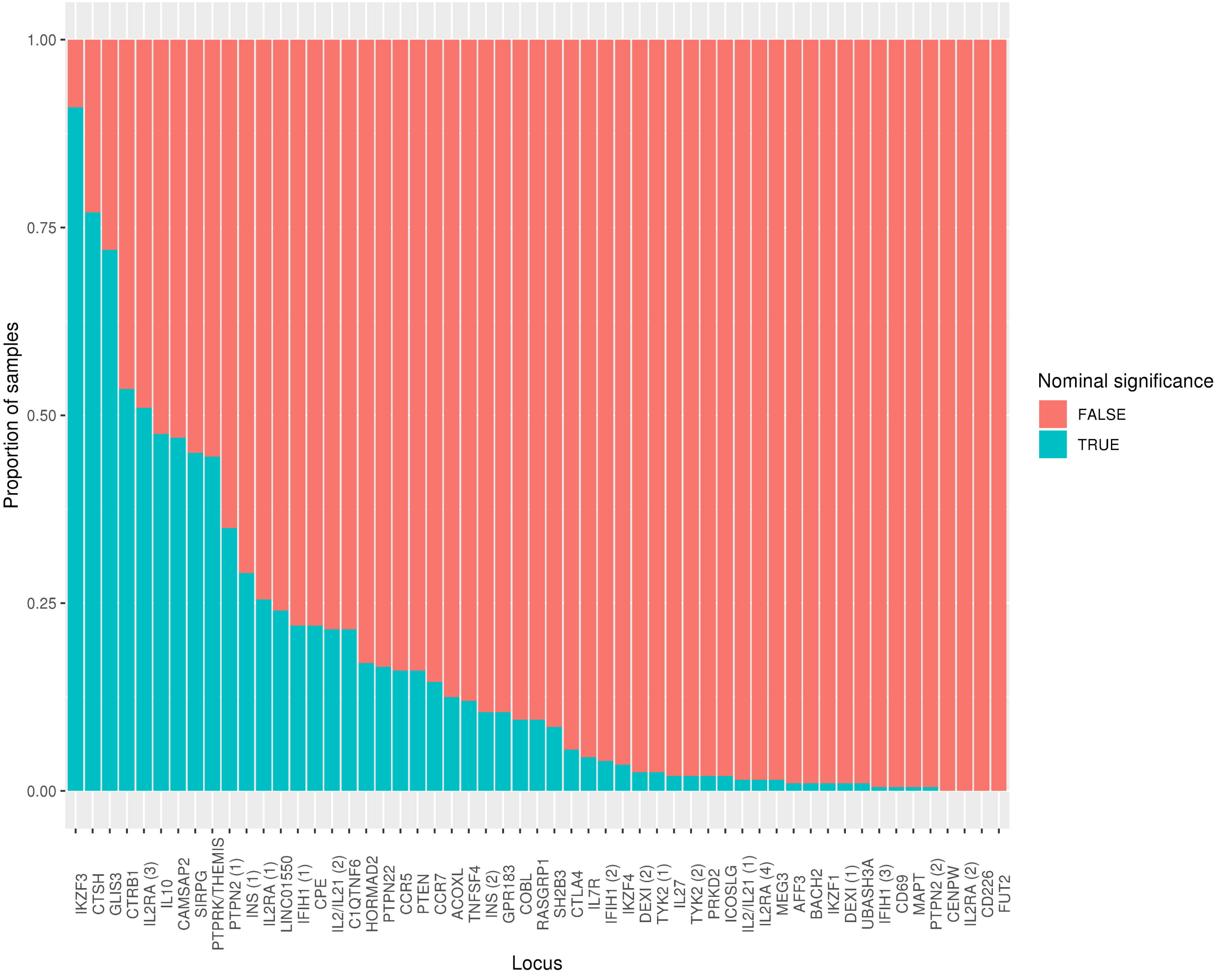

Locus
IKZF3
CTSH
GLIS3
CTRB1
IL2RA (3)
IL10
CAMSAP2
SIRPG
PTPRK/THEMIS
0
1
2
3
4
5
-log10 (p) Likelihood ratio test for heterogeneity between <7 and >13

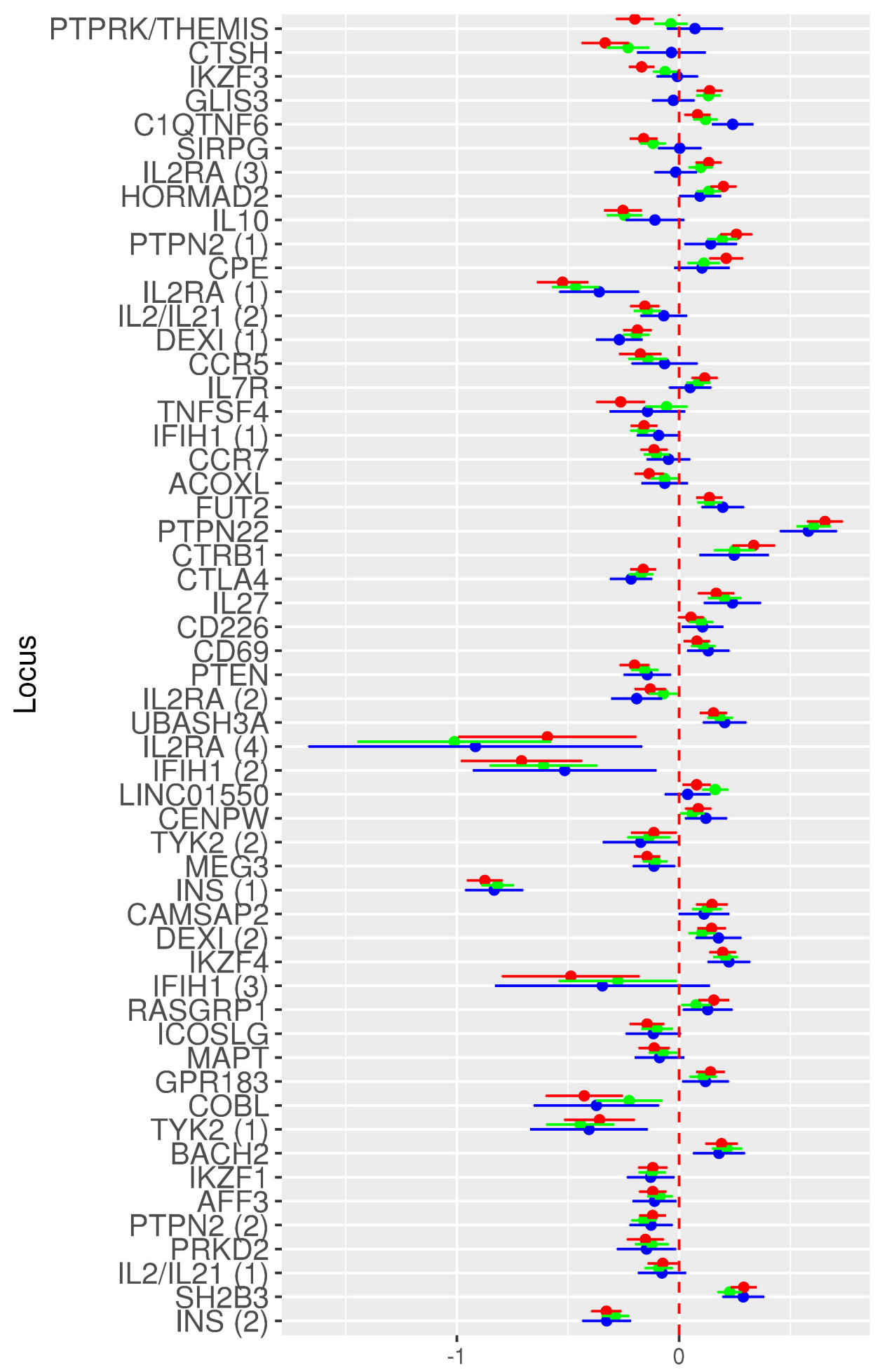


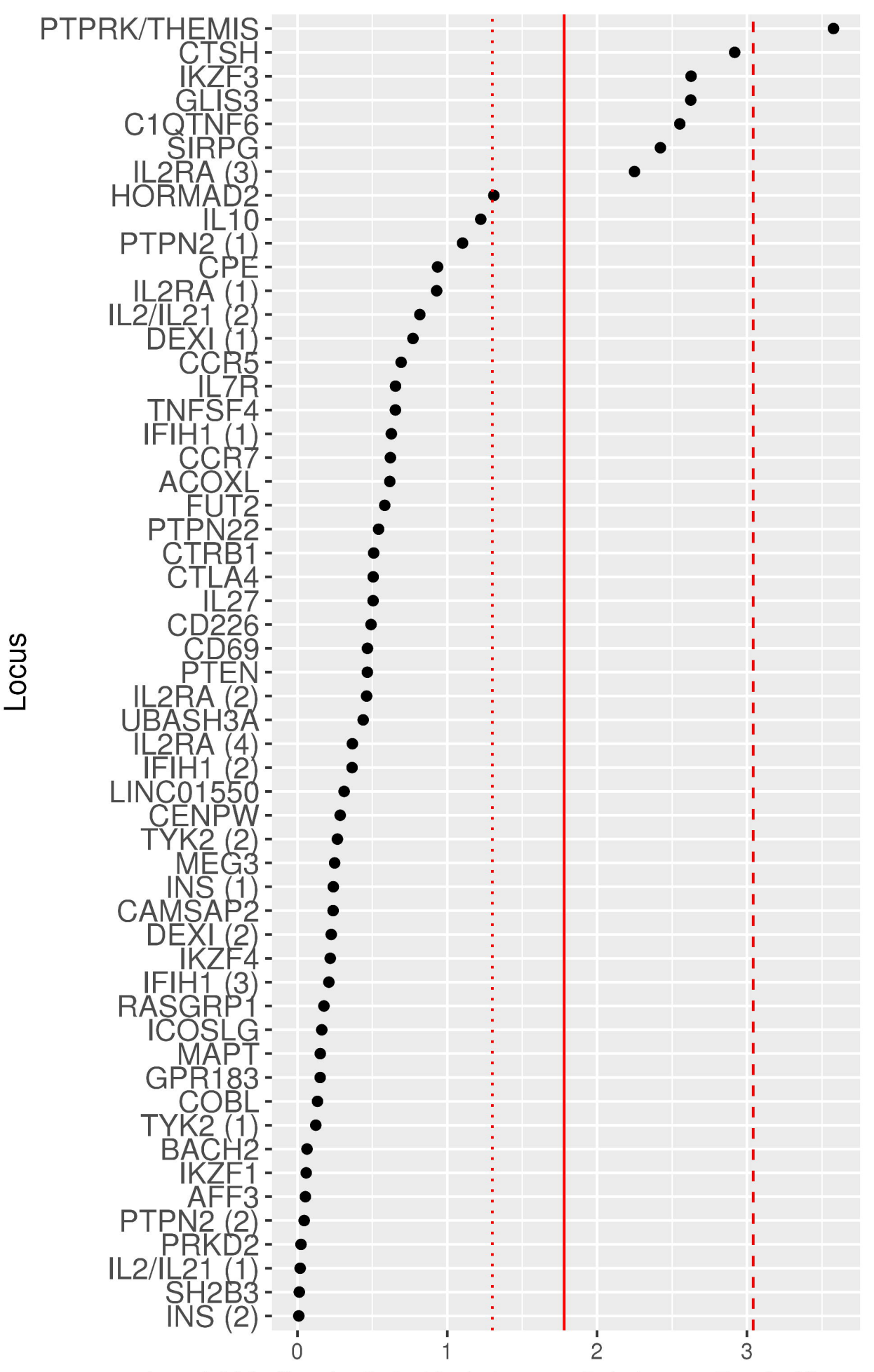

1.00
0.75
0.50
0.25
0.00
Proportion concordance on both chromosomes
N=1342
N=2685
N=4809
N=3809
A
B
DRB1
DQB1
Locus
HLA 4 digit concordance
FALSE
TRUE

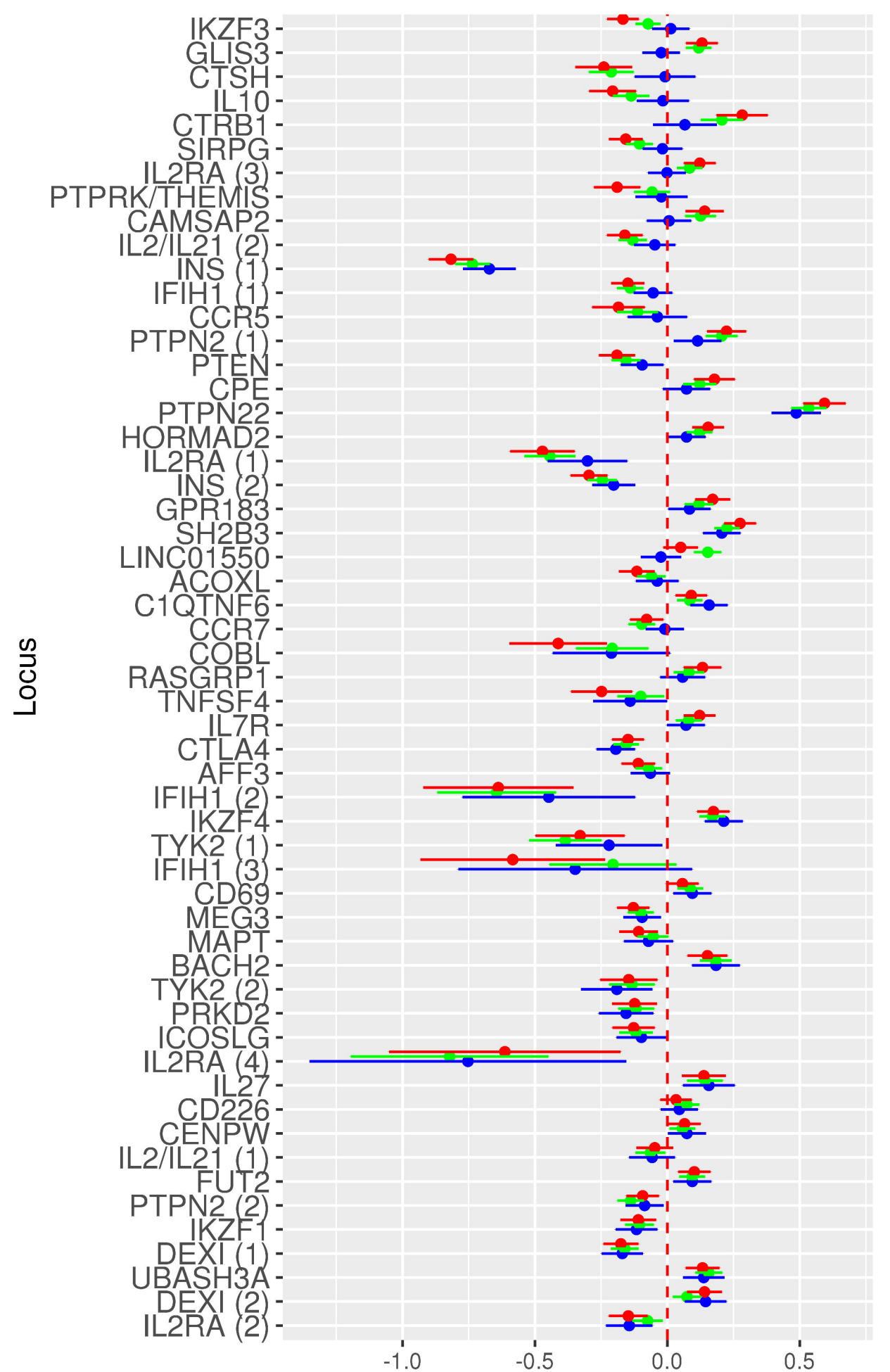

Locus
IKZF3
CTSH
GLIS3
CTRB1
IL2RA (3)
IL10
CAMSAP2
SIRPG
PTPRK/THEMIS
PTPN2 (1)
C1QTNF6
IL2/IL21 (2)
IL2RA (1)
CPE
INS (1)
LINC01550
HORMAD2
IFIH1 (1)
CCR7
PTEN
PTPN22
CCR5
ACOXL
GPR183
SH2B3
RASGRP1
TNFSF4
INS (2)
COBL
IFIH1 (2)
IKZF4
CTLA4
IL7R
TYK2 (1)
MEG3
CD69
TYK2 (2)
ICOSLG
IL2RA (4)
DEXI (2)
IL27
AFF3
MAPT
BACH2
IFIH1 (3)
PRKD2
IKZF1
CD226
PTPN2 (2)
DEXI (1)
CENPW
UBASH3A
FUT2
IL2/IL21 (1)
IL2RA (2)
-1.0
-0.5
0.0
0.5
T1D log-odds ratio for those diagnosed under 7 (red), 7-13 (green) and over 13 (blue)

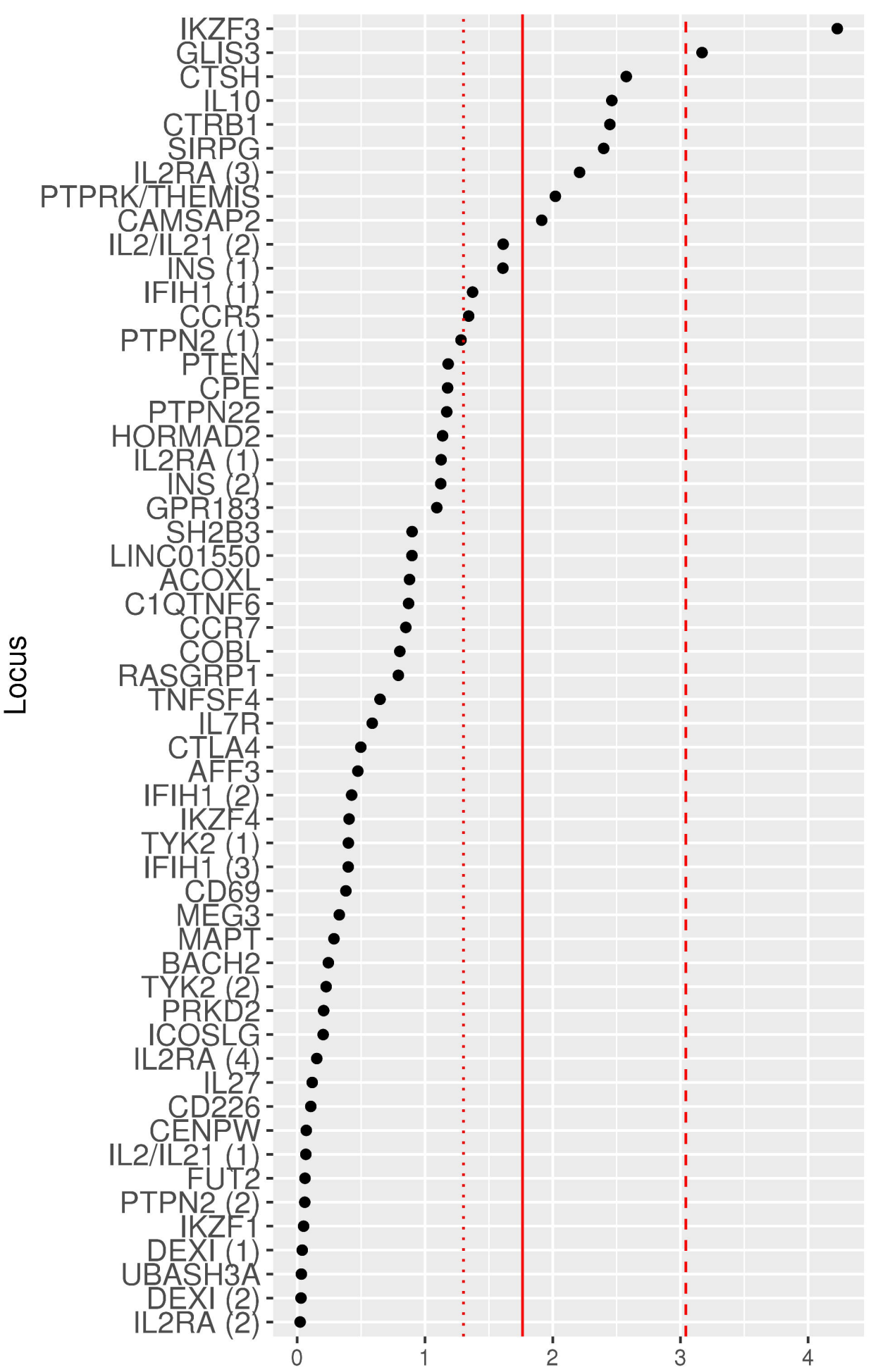

Locus
IKZF3
CTSH
GLIS3
CTRB1
IL2RA (3)
IL10
CAMSAP2
SIRPG
PTPRK/THEMIS
PTPN2 (1)
C1QTNF6
IL2/IL21 (2)
IL2RA (1)
CPE
INS (1)
LINC01550
HORMAD2
IFIH1 (1)
CCR7
PTEN
PTPN22
CCR5
ACOXL
GPR183
SH2B3
RASGRP1
TNFSF4
INS (2)
COBL
IFIH1 (2)
IKZF4
CTLA4
IL7R
TYK2 (1)
MEG3
CD69
TYK2 (2)
ICOSLG
IL2RA (4)
DEXI (2)
IL27
AFF3
MAPT
BACH2
IFIH1 (3)
PRKD2
IKZF1
CD226
PTPN2 (2)
DEXI (1)
CENPW
UBASH3A
FUT2
IL2/IL21 (1)
IL2RA (2)
0
1
2
3
4
-log10 (p) Likelihood ratio test for heterogeneity between <7 and >13

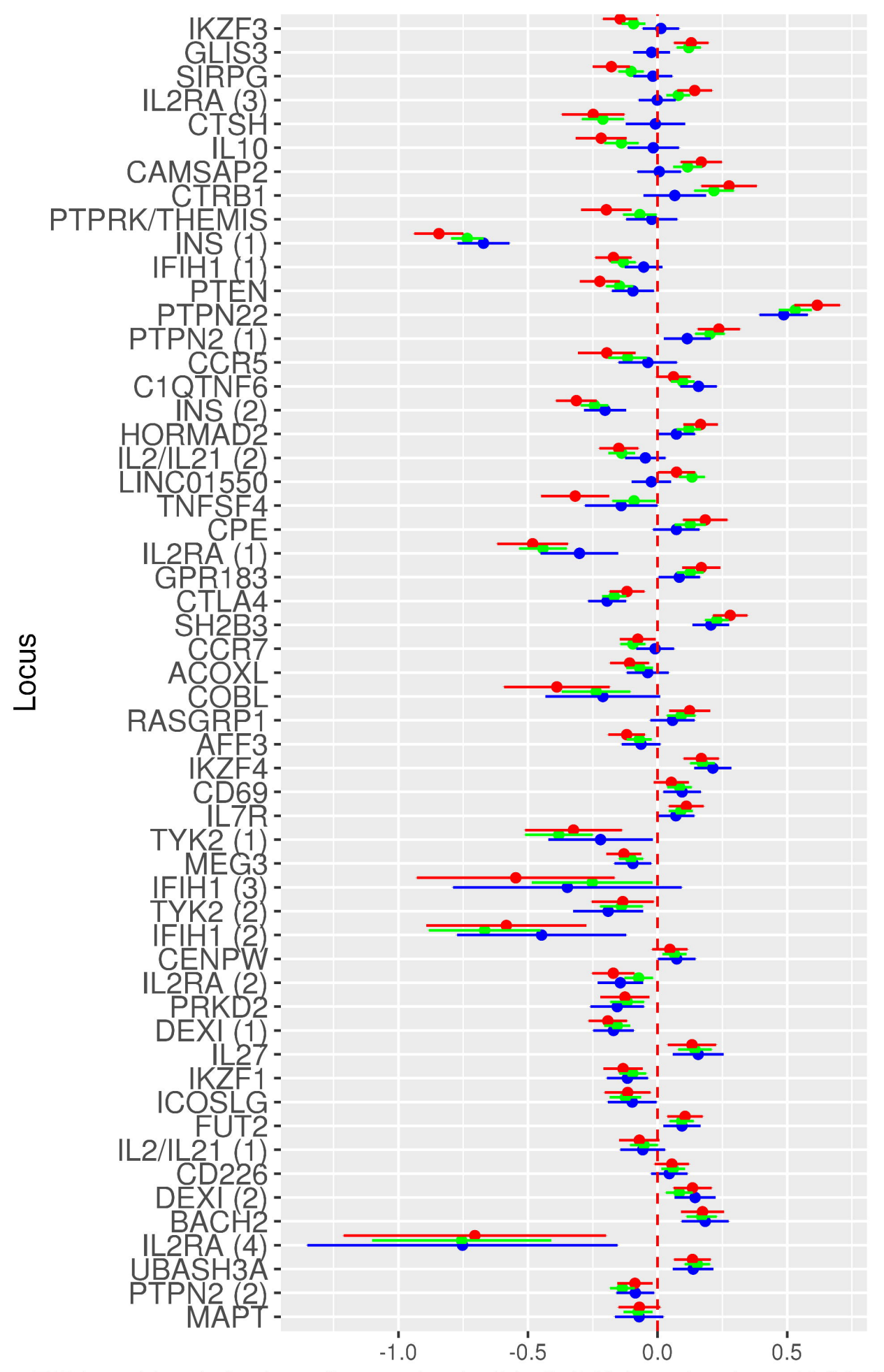

1.00
0.75
0.50
0.25
0.00
Proportion of samples
Nominal significance
FALSE
TRUE
IKZF3
CTSH
GLIS3
CTRB1
IL2RA (3)
IL10
CAMSAP2
SIRPG
PTPRK/THEMIS
PTPN2 (1)
INS (1)
IL2RA (1)
LINC01550
IFIH1 (1)
CPE
IL2/IL21 (2)
C1QTNF6
HORMAD2
PTPN22
CCR5
PTEN
CCR7
ACOXL
TNFSF4
INS (2)
GPR183
COBL
RASGRP1
SH2B3
CTLA4
IL7R
IFIH1 (2)
IKZF4
DEXI (2)
TYK2 (1)
IL27
TYK2 (2)
PRKD2
ICOSLG
IL2/IL21 (1)
IL2RA (4)
MEG3
AFF3
BACH2
IKZF1
DEXI (1)
UBASH3A
IFIH1 (3)
CD69
MAPT
PTPN2 (2)
CENPW
IL2RA (2)
CD226
FUT2
Locus

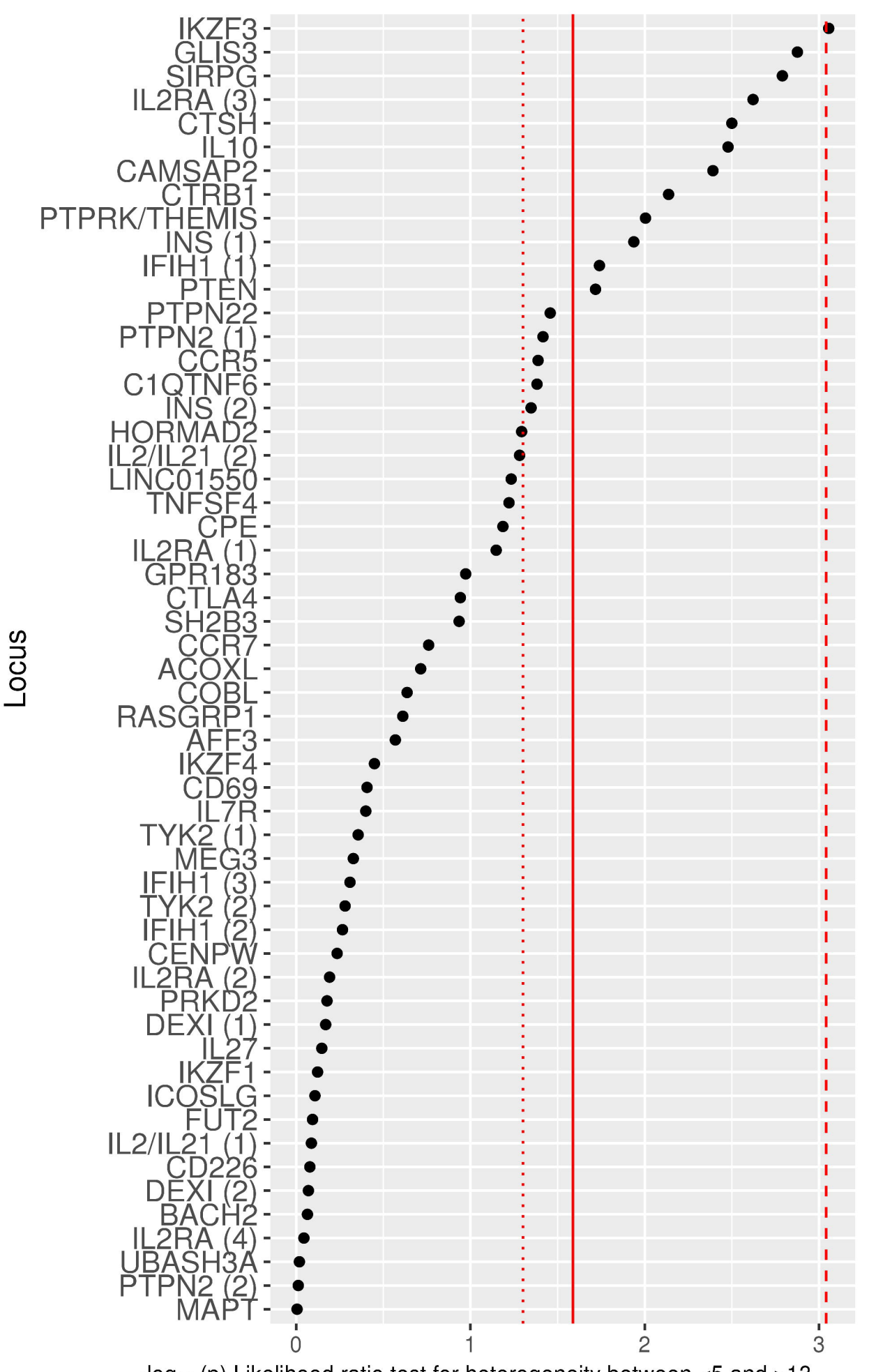

Locus
PTPRK/THEMIS
CTSH
IKZF3
GLIS3
C1QTNF6
SIRPG
IL2RA (3)
HORMAD2
IL10
PTPN2 (1)
CPE
IL2RA (1)
IL2/IL21 (2)
DEXI (1)
CCR5
IL7R
TNFSF4
IFIH1 (1)
CCR7
ACOXL
FUT2
PTPN22
CTRB1
CTLA4
IL27
CD226
CD69
PTEN
IL2RA (2)
UBASH3A
IL2RA (4)
IFIH1 (2)
LINC01550
CENPW
TYK2 (2)
MEG3
INS (1)
CAMSAP2
DEXI (2)
IKZF4
IFIH1 (3)
RASGRP1
ICOSLG
MAPT
GPR183
COBL
TYK2 (1)
BACH2
IKZF1
AFF3
PTPN2 (2)
PRKD2
IL2/IL21 (1)
SH2B3
INS (2)
-1
0
T1D log-odds ratio for those diagnosed under 7 (red), 7-13 (green) and over 13 (blue)


Locus
PTPRK/THEMIS
CTSH
IKZF3
GLIS3
C1QTNF6
SIRPG
IL2RA (3)
HORMAD2
IL10
PTPN2 (1)
CPE
IL2RA (1)
IL2/IL21 (2)
DEXI (1)
CCR5
IL7R
TNFSF4
IFIH1 (1)
CCR7
ACOXL
FUT2
PTPN22
CTRB1
CTLA4
IL27
CD226
CD69
PTEN
IL2RA (2)
UBASH3A
IL2RA (4)
IFIH1 (2)
LINC01550
CENPW
TYK2 (2)
MEG3
INS (1)
CAMSAP2
DEXI (2)
IKZF4
IFIH1 (3)
RASGRP1
ICOSLG
MAPT
GPR183
COBL
TYK2 (1)
BACH2
IKZF1
AFF3
PTPN2 (2)
PRKD2
IL2/IL21 (1)
SH2B3
INS (2)
0
1
2
3
-log10 (p) Likelihood ratio test for heterogeneity between <7 and >13


Locus
IKZF3
GLIS3
CTSH
IL10
CTRB1
SIRPG
IL2RA (3)
PTPRK/THEMIS
CAMSAP2
IL2/IL21 (2)
INS (1)
IFIH1 (1)
CCR5
PTPN2 (1)
PTEN
CPE
PTPN22
HORMAD2
IL2RA (1)
INS (2)
GPR183
SH2B3
LINC01550
ACOXL
C1QTNF6
CCR7
COBL
RASGRP1
TNFSF4
IL7R
CTLA4
AFF3
IFIH1 (2)
IKZF4
TYK2 (1)
IFIH1 (3)
CD69
MEG3
MAPT
BACH2
TYK2 (2)
PRKD2
ICOSLG
IL2RA (4)
IL27
CD226
CENPW
IL2/IL21 (1)
FUT2
PTPN2 (2)
IKZF1
DEXI (1)
UBASH3A
DEXI (2)
IL2RA (2)
-1.0
-0.5
0.0
0.5
T1D log-odds ratio for those diagnosed under 6 (red), 6-13 (green) and over 13 (blue)


Locus
IKZF3
GLIS3
CTSH
IL10
CTRB1
SIRPG
IL2RA (3)
PTPRK/THEMIS
CAMSAP2
IL2/IL21 (2)
INS (1)
IFIH1 (1)
CCR5
PTPN2 (1)
PTEN
CPE
PTPN22
HORMAD2
IL2RA (1)
INS (2)
GPR183
SH2B3
LINC01550
ACOXL
C1QTNF6
CCR7
COBL
RASGRP1
TNFSF4
IL7R
CTLA4
AFF3
IFIH1 (2)
IKZF4
TYK2 (1)
IFIH1 (3)
CD69
MEG3
MAPT
BACH2
TYK2 (2)
PRKD2
ICOSLG
IL2RA (4)
IL27
CD226
CENPW
IL2/IL21 (1)
FUT2
PTPN2 (2)
IKZF1
DEXI (1)
UBASH3A
DEXI (2)
IL2RA (2)
0
1
2
3
4
-log10 (p) Likelihood ratio test for heterogeneity between <6 and >13


Locus
IKZF3
GLIS3
SIRPG
IL2RA (3)
CTSH
IL10
CAMSAP2
CTRB1
PTPRK/THEMIS
INS (1)
IFIH1 (1)
PTEN
PTPN22
PTPN2 (1)
CCR5
C1QTNF6
INS (2)
HORMAD2
IL2/IL21 (2)
LINC01550
TNFSF4
CPE
IL2RA (1)
GPR183
CTLA4
SH2B3
CCR7
ACOXL
COBL
RASGRP1
AFF3
IKZF4
CD69
IL7R
TYK2 (1)
MEG3
IFIH1 (3)
TYK2 (2)
IFIH1 (2)
CENPW
IL2RA (2)
PRKD2
DEXI (1)
IL27
IKZF1
ICOSLG
FUT2
IL2/IL21 (1)
CD226
DEXI (2)
BACH2
IL2RA (4)
UBASH3A
PTPN2 (2)
MAPT
-1.0
-0.5
0.0
0.5
T1D log-odds ratio for those diagnosed under 5 (red), 5-13 (green) and over 13 (blue)


Locus
IKZF3
GLIS3
SIRPG
IL2RA (3)
CTSH
IL10
CAMSAP2
CTRB1
PTPRK/THEMIS
INS (1)
IFIH1 (1)
PTEN
PTPN22
PTPN2 (1)
CCR5
C1QTNF6
INS (2)
HORMAD2
IL2/IL21 (2)
LINC01550
TNFSF4
CPE
IL2RA (1)
GPR183
CTLA4
SH2B3
CCR7
ACOXL
COBL
RASGRP1
AFF3
IKZF4
CD69
IL7R
TYK2 (1)
MEG3
IFIH1 (3)
TYK2 (2)
IFIH1 (2)
CENPW
IL2RA (2)
PRKD2
DEXI (1)
IL27
IKZF1
ICOSLG
FUT2
IL2/IL21 (1)
CD226
DEXI (2)
BACH2
IL2RA (4)
UBASH3A
PTPN2 (2)
MAPT
0
1
2
3
-log10 (p) Likelihood ratio test for heterogeneity between <5 and >13

IKZF3

0.5

0.4

Minor Allele Frequency

Controls

1

2

3

4

5

6

7

8

9

10

11

12

13

14

15

16-20

20-25

25-30

30-35

>35

Age-at-diagnosis

CTSH

Minor Allele Frequency

0.25

0.20

0.15

0.10

0.05

Controls

1

2

3

4

5

6

7

8

9

10

11

12

13

14

15

16-20

20-25

25-30

30-35

>35

Age-at-diagnosis

GLIS3

Minor Allele Frequency

0.55
0.50
0.45
0.40
0.35
0.30

Controls
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16-20
20-25
25-30
30-35
>35

Age-at-diagnosis

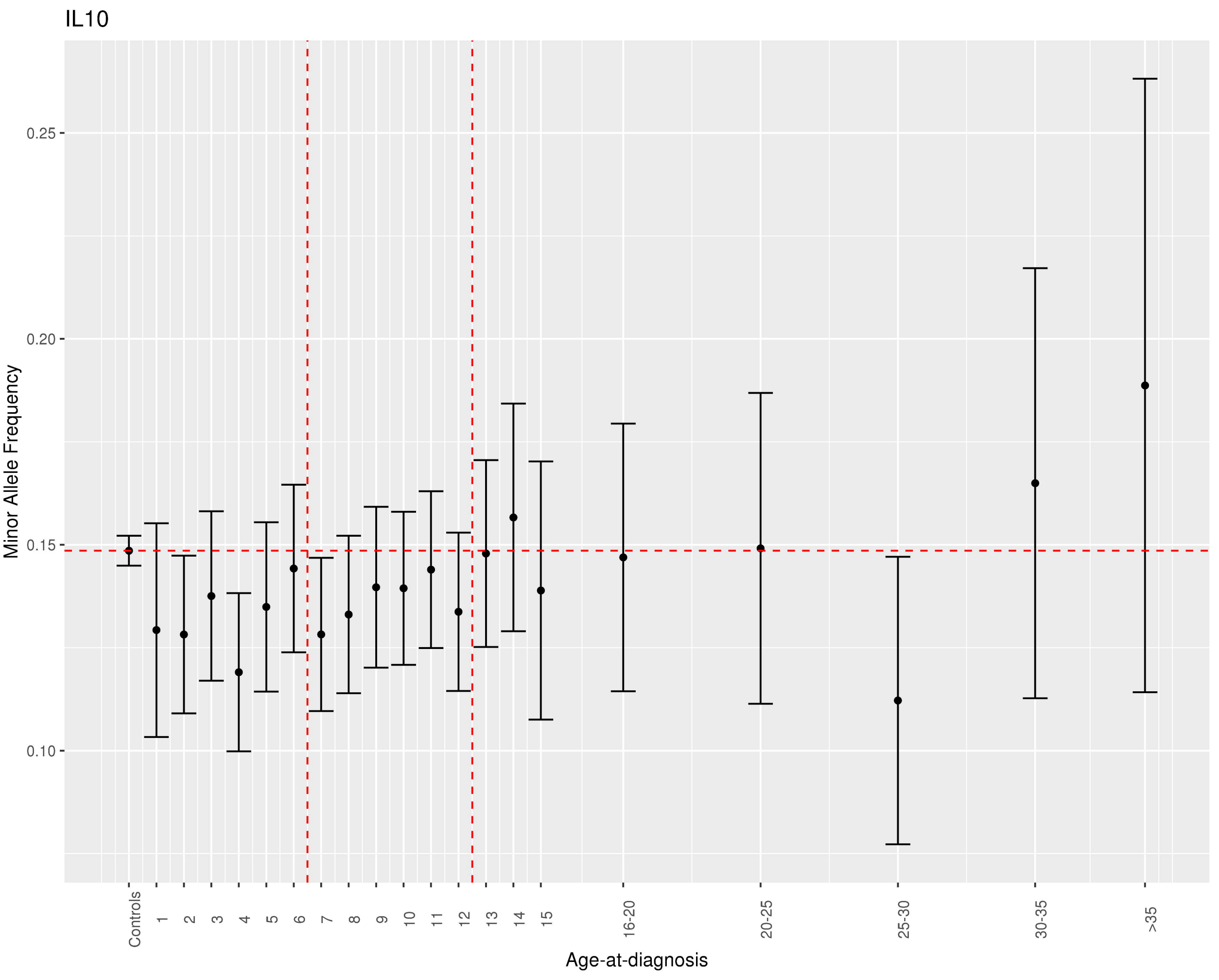
IL2RA (3)

Minor Allele Frequency

0.55
0.50
0.45
0.40
0.35

Controls
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16-20
20-25
25-30
30-35
>35

Age-at-diagnosis


IL10
Minor Allele Frequency
0.25
0.20
0.15
0.10
Controls
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16-20
20-25
25-30
30-35
>35
Age-at-diagnosis

SIRPG

Minor Allele Frequency

0.40
0.35
0.30
0.25
0.20

Controls
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16-20
20-25
25-30
30-35
>35

Age-at-diagnosis

PTPRK/THEMIS

Minor Allele Frequency

0.20

0.15

0.10

Controls

1

2

3

4

5

6

7

8

9

10

11

12

13

14

15

16-20

20-25

25-30

30-35

>35

Age-at-diagnosis