

Abstract

 Loss-of-function variants in MDA5, a key sensor of double-stranded RNA from viruses and retroelements, have been associated with protection from type 1 diabetes (T1D) in genome- wide association studies (GWAS). MDA5 loss-of-function variants have also been reported to increase the risk of inflammatory bowel disease (IBD). Whether these associations are linked or extend to other diseases remains unclear. Here, fine-mapping analysis of four large GWAS datasets shows that T1D-protective loss-of-function MDA5 variants also protect against psoriasis and hypothyroidism, while increasing the risk of IBD. The degree of autoimmune 23 protection and IBD risk were linearly proportional. The magnitudes of the odds ratios for autoimmune protection and IBD risk were larger for rare MDA5 variants than for common variants, which were differentially expressed in different geographic populations. Our analysis suggests MDA5 genetic variants offer a direct fitness trade-off between viral clearance and autoimmune tissue damage.

Introduction

 Innate immune responses must be sensitive enough to detect infection yet specific enough to avoid activation by cellular components. A key innate immune receptor for cytosolic double-32 stranded RNA (dsRNA), a potent signature of viral infection, is MDA5¹⁻⁵. Recognition of dsRNA 33 by MDA5 induces a potent IFN- β response^{2,3}. The ATPase activity and dsRNA binding cooperativity of MDA5 confer the necessary sensitivity and specificity of MDA5-dsRNA 35 recognition⁴⁻⁶. Mutations in the gene encoding MDA5, *IFIH1*, can perturb the balance between sensitive and specific dsRNA recognition. *IFIH1* is a hotspot for natural variants with clinical associations. Approximately 40 gain-of-function missense variants have been associated with 38 interferonopathies⁷⁻⁹. These variants promote formation of MDA5 signaling complexes, including on endogenous dsRNAs, by either increasing the RNA binding affinity of MDA5 or inhibiting its

 40 ATP-dependent proofreading activity $5,9,10$. Loss-of-function MDA5 variants cause recurrent 11 infections¹¹, and have been reported to contribute to inflammatory bowel disease (IBD), 42 including ulcerative colitis and Crohn's disease^{12,13}. GWAS data have also associated loss-of-43 function MDA5 variants with reduced risk of developing certain autoimmune diseases, most 44 notably T1D¹⁴⁻¹⁹. Specifically, variants E627^{*} (rs35744605), R843H (rs3747517), I923V (rs35667974), and T946A (rs1990760) have been identified as T1D-protective^{14-18,20,21}. The 46 E627* and I923V variants are rare, with allele frequencies of 1-2%, while R843H and T946A are 47 common. These common variants are differentially expressed in different geographic 48 populations. The T946A variant is present in 70-80% of Africans and Asians but only 30-50% of 49 Caucasians^{15-18,20}. Similarly, the R843H variant is found in 70% of Asians but only 30-40% of 50 Caucasians and Africans^{15-18,20}. Most human *IFIH1* reference sequences contain the alleles that 51 are most common in Asians (Ala946/His843). These coding variants inhibit the formation of 52 MDA5-dsRNA signaling complexes^{18,22} by reducing the RNA binding affinity of MDA5 or 53 hyperactivating its ATPase activity²³. Other loss-of-function variants are splice donor variants 54 that reduce splicing efficiency and hence mRNA levels¹⁹.

55 A robust clinical link has emerged between T1D onset and recent infection with RNA 56 viruses, in particular coxsackieviruses and other enteroviruses 24 . T1D patients have more 57 frequent and persistent enterovirus infections, which precede the appearance of prediabetic 58 markers, including autoantibodies²⁵. MDA5 recognizes RNA from *Picornaviridae*, including 59 enteroviruses^{1,26}, which have evolved mechanisms to suppress IFN- β transcription^{27,28}. MDA5-60 induced inflammation in the pancreas following rotavirus infection contributes to autoimmune 61 destruction of pancreatic β -cells²⁹. Therefore, a plausible hypothesis is that MDA5-dependent 62 inflammation following viral infection can trigger autoimmune β -cell killing.

 Gastrointestinal viral infection has also been linked to IBD onset³⁰. Intestinal cells from IBD patients harboring loss-of-function MDA5 variants had elevated viral loads and were compromised in their ability to maintain epithelial barrier integrity upon further exposure to the enteric virome¹³. Hence, the frequency and clinical phenotypes of MDA5 loss-of-function 67 variants in IBD patients suggest that MDA5 deficiency contributes to the induction of IBD^{12,13}. μ due in part to increased exposure and susceptibility to viral infection^{11,13}.

Results and Discussion

 Loss-of-function MDA5 variants have been associated with recurrent infection¹¹, T1D 71 protection¹⁴⁻¹⁹, and IBD risk^{12,13}. However, whether these associations are linked, or whether they apply to autoimmune disease more broadly, has not been explored. We addressed these questions by utilizing larger, more recent GWAS studies to conduct a broader, more sensitive survey of immune-mediated disease associations involving MDA5 variants. For our analysis, we selected GWAS datasets to maximize the likelihood of accurate fine mapping – balancing the need for large sample size, minimizing the use of imputed genotypes, and using ancestries that best matched the reference panel used to derive linkage disequilibrium (LD) estimates³¹⁻³⁴ (see **Methods**). By reviewing associations with a reported lead T1D-associated variant, I923V (79) (rs35667974)¹⁵ we identified psoriasis, hypothyroidism, Crohn's disease, and ulcerative colitis 80 as additional associated diseases. Fine-mapping analysis using GWAS summary data for these five diseases identified four variants likely to be causally associated with a subset of the diseases, with considerable overlap of variants between diseases (**Fig. 1, Table 1**). Two of these variants were coding variants, I923V and T946A, and two were splice donor variants, 84 rs35337543 and rs35732034. All four variants were previously associated with T1D protection¹⁴⁻ 19

 Given the overlaps between diseases, and the small *P* values observed for fine-mapped variants in one disease when considering another, we called associated variants by considering marginal association at any of these variants or other known loss-of-function variants not in LD with any of the index variants using a stringent Bonferroni correction. This identified associations of a fifth variant, E627* (rs35744605), with psoriasis, and hypothyroidism (**Fig. 1**). While T946A is common, the other variants are rare (allele frequency below 3%). Fine-mapping analysis did not find evidence for causality for variants E627* and R843H (**Table 1**). However, for E627* the lack of evidence of causality is likely due to its low allele frequency (0.7%), which limits power in genetics studies. In contrast, association of R843H with T1D protection is likely 95 attributable to LD with the co-occurring causal T946A allele (r^2 = 0.60), as previously 96 reported^{15,16,18}. Supporting these conclusions, our accompanying biochemical and structural study of T1D-protective MDA5 variants shows that the E627* variant lacks signaling activity due to a loss of RNA binding affinity, whereas the R843H substitution had no effect on the structural, 99 biochemical or signaling activities of MDA 5^{23} .

 Comparing associations across diseases, we found that all five associated variants offered protection against T1D, psoriasis and hypothyroidism, but increased the risk of ulcerative colitis and Crohn's disease (**Fig. 2a**). Remarkably, a strict correlation was observed between protection and risk (**Fig. 2b**). A statistical test of proportionality (see **Methods**) concluded that changes in risk were linearly proportional across variants (**Fig. 2c**). Additionally, the magnitudes of the odds ratios were larger for the rare variants than for the common variants. Thus, the rarest variant, I923V, conferred the greatest degree of T1D protection and the greatest risk of IBD, whereas the most common variant, T946A, was associated with the least T1D protection and IBD risk (**Fig. 2a-b**). Notably, the T946A variant, while common in all populations, is the major allele in Asian and African populations but the minor allele in Caucasians.

- GCST006000/GCST005527/harmonised/23143594-GCST005527-EFO_0000676.h.tsv.gz
- (Meta analysis); Hypothyroidism: https://pan-ukb-us-east-
- 138 1.s3.amazonaws.com/sumstats_flat_files/categorical-20002-both_sexes-1453.tsv.bgz (genome
- wide SNP chip). These studies were selected to balance the need for large studies and those
- most likely to give accurate results in fine mapping i.e. those which minimized imputation and
- most closely matched the ancestry of our reference LD panel.
- Alleles were aligned to the UK Biobank as a common reference. Reference LD matrices
- were estimated from 40,000 European subjects from UK Biobank. Fine mapping was performed
- 144 with a variant of the Sum of Single Effects (SuSiE)³⁵ model, as implemented in susieR³⁶.
- 145 rs72871627 was selected as a tag (r^2 = 0.99) of the splice donor variant rs35337543
- because rs72871627 was available in all GWAS datasets whereas rs35337543 was only
- available in the T1D GWAS dataset.
- Tests of proportional effects between pairs of diseases were performed across all variants in
- 149 Table X using colocPropTest [https://cran.r-project.org/package=colocPropTest] which
- 150 implements the test of proportionality as previously described .

Data Availability

- This study is a meta-analysis of previously published data. All code to reproduce this analysis is
- available at Zenodo.org [https://doi.org/10.5281/zenodo.12771481].

Competing Interests

- Y.M. is a consultant for Related Sciences LLC. The remaining authors declare no competing
- interests.
- **Author Contributions**

- Conceptualization, C.E.W., R.S. and Y.M.; Methodology, C.E.W.; Formal Analysis, C.E.W. and
- Y.M.; Funding Acquisition, Y.M.; Project Administration, Y.M.; Software, C.E.W.; Supervision,
- Y.M.; Writing Original Draft, Y.M.; Writing Review & Editing, C.E.W., R.S. and Y.M.

Acknowledgements

- We thank Louise Modis, Tetsuo Hasegawa, and the Modis group for useful discussions. This
- work was supported by the Wellcome Trust [217191/Z/19/Z to Y.M.].

References

1. Kato H*, et al.* Differential roles of MDA5 and RIG-I helicases in the recognition of RNA

viruses. *Nature* **441**, 101-105 (2006).

- 2. Peisley A*, et al.* Cooperative assembly and dynamic disassembly of MDA5 filaments for
- viral dsRNA recognition. *Proc Natl Acad Sci U S A* **108**, 21010-21015 (2011).
- 3. Berke IC, Modis Y. MDA5 cooperatively forms dimers and ATP-sensitive filaments upon

binding double-stranded RNA. *EMBO J* **31**, 1714-1726 (2012).

- 4. Yu Q, Qu K, Modis Y. Cryo-EM Structures of MDA5-dsRNA Filaments at Different Stages of ATP Hydrolysis. *Mol Cell* **72**, 999-1012 e1016 (2018).
- 5. Yu Q, Herrero Del Valle A, Singh R, Modis Y. MDA5 disease variant M854K prevents ATP-
- dependent structural discrimination of viral and cellular RNA. *Nat Commun* **12**, 6668 (2021).
- 6. Singh R*, et al.* Contrasting functions of ATP hydrolysis by MDA5 and LGP2 in viral RNA
- sensing. *J Biol Chem* **300**, 105711 (2024).
- 7. Rodero MP, Crow YJ. Type I interferon-mediated monogenic autoinflammation: The type I
- interferonopathies, a conceptual overview. *J Exp Med* **213**, 2527-2538 (2016).
- 8. Rutsch F*, et al.* A specific IFIH1 gain-of-function mutation causes Singleton-Merten
- syndrome. *Am J Hum Genet* **96**, 275-282 (2015).

9. Rice GI*, et al.* Genetic and phenotypic spectrum associated with IFIH1 gain-of-function.

Hum Mutat **41**, 837-849 (2020).

- 10. Garau J*, et al.* Molecular Genetics and Interferon Signature in the Italian Aicardi Goutieres
- Syndrome Cohort: Report of 12 New Cases and Literature Review. *J Clin Med* **8**, 750
- (2019).
- 11. Lamborn IT*, et al.* Recurrent rhinovirus infections in a child with inherited MDA5 deficiency.
- *J Exp Med* **214**, 1949-1972 (2017).
- 12. Cananzi M*, et al.* IFIH1 loss-of-function variants contribute to very early-onset inflammatory

bowel disease. *Hum Genet* **140**, 1299-1312 (2021).

- 13. Adiliaghdam F*, et al.* Human enteric viruses autonomously shape inflammatory bowel
- disease phenotype through divergent innate immunomodulation. *Sci Immunol* **7**, eabn6660 (2022).
- 14. Smyth DJ*, et al.* A genome-wide association study of nonsynonymous SNPs identifies a
- type 1 diabetes locus in the interferon-induced helicase (IFIH1) region. *Nat Genet* **38**, 617- 619 (2006).
- 15. Nejentsev S, Walker N, Riches D, Egholm M, Todd JA. Rare variants of IFIH1, a gene
- implicated in antiviral responses, protect against type 1 diabetes. *Science* **324**, 387-389 (2009).
- 16. Liu S*, et al.* IFIH1 polymorphisms are significantly associated with type 1 diabetes and
- IFIH1 gene expression in peripheral blood mononuclear cells. *Hum Mol Genet* **18**, 358-365 (2009).
- 17. Vasseur E*, et al.* The selective footprints of viral pressures at the human RIG-I-like receptor family. *Hum Mol Genet* **20**, 4462-4474 (2011).
- 18. Gorman JA*, et al.* The A946T variant of the RNA sensor IFIH1 mediates an interferon
- program that limits viral infection but increases the risk for autoimmunity. *Nat Immunol* **18**,
- 744-752 (2017).

- 19. Downes K*, et al.* Reduced expression of IFIH1 is protective for type 1 diabetes. *PLoS One* **5**, (2010).
- 20. de Azevedo Silva J*, et al.* Meta-analysis of STAT4 and IFIH1 polymorphisms in type 1
- diabetes mellitus patients with autoimmune polyglandular syndrome type III. *Genet Mol Res*
- **14**, 17730-17738 (2015).
- 21. Jermendy A, Szatmari I, Korner A, Szabo AJ, Toth-Heyn P, Hermann R. Association
- between interferon-induced helicase (IFIH1) rs1990760 polymorphism and seasonal
- variation in the onset of type 1 diabetes mellitus. *Pediatr Diabetes* **19**, 300-304 (2018).
- 22. Shigemoto T, Kageyama M, Hirai R, Zheng J, Yoneyama M, Fujita T. Identification of loss
- of function mutations in human genes encoding RIG-I and MDA5: implications for
- resistance to type I diabetes. *J Biol Chem* **284**, 13348-13354 (2009).
- 23. Singh R*, et al.* Molecular basis of pathogenesis and protection from type 1 diabetes by loss of function MDA5 variants. *bioRxiv*, (2024).
- 24. Yeung WC, Rawlinson WD, Craig ME. Enterovirus infection and type 1 diabetes mellitus:
- systematic review and meta-analysis of observational molecular studies. *BMJ* **342**, d35
- (2011).
- 25. Hyoty H, Taylor KW. The role of viruses in human diabetes. *Diabetologia* **45**, 1353-1361 (2002).
- 26. Wang JP, Cerny A, Asher DR, Kurt-Jones EA, Bronson RT, Finberg RW. MDA5 and MAVS
- mediate type I interferon responses to coxsackie B virus. *J Virol* **84**, 254-260 (2010).
- 27. Visser LJ*, et al.* Essential Role of Enterovirus 2A Protease in Counteracting Stress Granule
- Formation and the Induction of Type I Interferon. *J Virol* **93**, (2019).
- 28. Hato SV*, et al.* The mengovirus leader protein blocks interferon-alpha/beta gene
- transcription and inhibits activation of interferon regulatory factor 3. *Cell Microbiol* **9**, 2921-
- 2930 (2007).

- 29. Dou Y, Yim HC, Kirkwood CD, Williams BR, Sadler AJ. The innate immune receptor MDA5 limits rotavirus infection but promotes cell death and pancreatic inflammation. *EMBO J* **36**, 2742-2757 (2017).
- 30. Axelrad JE*, et al.* Gastrointestinal Infection Increases Odds of Inflammatory Bowel Disease
- in a Nationwide Case-Control Study. *Clin Gastroenterol Hepatol* **17**, 1311-1322 e1317
- (2019).
- 31. Robertson CC*, et al.* Fine-mapping, trans-ancestral and genomic analyses identify causal variants, cells, genes and drug targets for type 1 diabetes. *Nat Genet* **53**, 962-971 (2021).
- 32. Liu JZ*, et al.* Association analyses identify 38 susceptibility loci for inflammatory bowel
- disease and highlight shared genetic risk across populations. *Nat Genet* **47**, 979-986
- (2015).
- 33. Tsoi LC*, et al.* Large scale meta-analysis characterizes genetic architecture for common psoriasis associated variants. *Nat Commun* **8**, 15382 (2017).
- 34. Karczewski KJ*, et al.* Pan-UK Biobank GWAS improves discovery, analysis of genetic
- architecture, and resolution into ancestry-enriched effects. *bioRxiv*,
- 2024.2003.2013.24303864 (2024).
- 35. Wang G, Sarkar A, Carbonetto P, Stephens M. A simple new approach to variable selection
- in regression, with application to genetic fine mapping. *J R Stat Soc Series B Stat Methodol*
- **82**, 1273-1300 (2020).
- 36. Zou Y, Carbonetto P, Wang G, Stephens M. Fine-mapping from summary data with the "Sum of Single Effects" model. *PLoS genetics* **18**, e1010299 (2022).
- 37. Wallace C. Statistical testing of shared genetic control for potentially related traits. *Genet Epidemiol* **37**, 802-813 (2013).
-
-

258

259 **Fig. 1. Manhattan plots of fine-mapped variants for five diseases.** Colored circles indicate 260 the fine mapping sets identified, and colored text labels the most likely causal SNP in each set, 261 selected due its known function on MDA5. Known loss-of-function SNPs that were not in any 262 fine mapping set are marked with "x" and a black label. Lower right, minor allele frequency 263 (MAF) and function of fine-mapped SNPs.

It is made available under a CC-BY 4.0 International license. **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. medRxiv preprint doi: [https://doi.org/10.1101/2024.10.04.24314884;](https://doi.org/10.1101/2024.10.04.24314884) this version posted October 4, 2024. The copyright holder for this preprint

264 265 **Figure 2. GWAS meta-analysis of loss-of-function MDA5 variants. (a)** Effect sizes expressed in terms 266 of the odd ratio for selected MDA5 loss-of-function variants. 4 Because rs35337543 was only available in 267 two GWAS datasets, rs72871627 (r^2 = 0.973) was used as a tag for rs35337543. EAF, effect allele 267 two GWAS datasets, rs72871627 (r^2 = 0.973) was used as a tag for rs35337543. EAF, effect allele 268 frequency. Error bars, SEM. ns, P > 0.05; *, P < 0.05; **, P < 0.01; ***, P < 0.001; ****, P < 5x10⁻⁸ 268 frequency. Error bars, SEM. ns, P > 0.05; *, P < 0.05; **, P < 0.01; ***, P < 0.001; ****, P < 5×10⁻⁸ (genome-wide significance), where P is the Bonferroni-adjusted P value from marginal association 269 (genome-wide significance), where P is the Bonferroni-adjusted P value from marginal association 270 analysis. CV, likely causal variant present in a credible set in fine-mapping analysis. R843H was included
271 in fine-mapping analysis but excluded from marginal association analysis. See Table 1. (b) Odds ratio 271 in fine-mapping analysis but excluded from marginal association analysis. See **Table 1**. **(b)** Odds ratio 272 plot for variants and diseases in *(a)*. (c) Estimated effect sizes in hypothyroidism compared to the other
273 diseases. Hypothyroidism was selected because it was the largest dataset for which all SNPs were 273 diseases. Hypothyroidism was selected because it was the largest dataset for which all SNPs were 274 available. Dashed lines, identity and negative identity lines. available. Dashed lines, identity and negative identity lines. 275

276 **Table 1. Marginal association and fine-mapping analysis results.**

277 EAF, effect allele frequency; OR, odds ratio.
278 * rs72871627 was used as a tag of rs353375

278 *rs72871627 was used as a tag of rs35337543 because the latter was not available in all GWAS datasets.

 280 ^tCausality was inferred from presence of the SNP in a credible set in fine mapping analysis.

281 Dashed lines indicate that the SNPs above and below the line are in LD with each other.