


RESEARCH

Open Access



Saturation-scale functional evidence supports clinical variant interpretation in Lynch syndrome

Anthony Scott^{1,2}, Felicia Hernandez³, Adam Chamberlin³, Cathy Smith^{1,4}, Rachid Karam^{3,4} and Jacob O. Kitzman^{1,4*} 

*Correspondence:
kitzmanj@umich.edu

¹ Department of Human Genetics, University of Michigan Medical School, Ann Arbor, MI 48109, USA

² Division of Genetic Medicine, Department of Internal Medicine, University of Michigan Medical School, Ann Arbor, MI 48109, USA

³ Ambry Genetics, Aliso Viejo, CA 92656, USA

⁴ Department of Computational Medicine and Bioinformatics, University of Michigan Medical School, Ann Arbor, MI 48109, USA

Abstract

Background: Lynch syndrome (LS) is a cancer predisposition syndrome affecting more than 1 in every 300 individuals worldwide. Clinical genetic testing for LS can be life-saving but is complicated by the heavy burden of variants of uncertain significance (VUS), especially missense changes.

Result: To address this challenge, we leverage a multiplexed analysis of variant effect (MAVE) map covering >94% of the 17,746 possible missense variants in the key LS gene *MSH2*. To establish this map's utility in large-scale variant reclassification, we overlay it on clinical databases of >15,000 individuals with LS gene variants uncovered during clinical genetic testing. We validate these functional measurements in a cohort of individuals with paired tumor-normal test results and find that MAVE-based function scores agree with the clinical interpretation for every one of the *MSH2* missense variants with an available classification. We use these scores to attempt reclassification for 682 unique missense VUS, among which 34 scored as deleterious by our function map, in line with previously published rates for other cancer predisposition genes. Combining functional data and other evidence, ten missense VUS are reclassified as pathogenic/likely pathogenic, and another 497 could be moved to benign/likely benign. Finally, we apply these functional scores to paired tumor-normal genetic tests and identify a subset of patients with biallelic somatic loss of function, reflecting a sporadic Lynch-like Syndrome with distinct implications for treatment and relatives' risk.

Conclusion: This study demonstrates how high-throughput functional assays can empower scalable VUS resolution and prospectively generate strong evidence for variant classification.

Background

Identification of a pathogenic variant in a familial cancer risk gene can inform treatment and prevention strategies for patients and their blood relatives. As a prominent example of a common and broadly screened cancer risk syndrome, Lynch syndrome (LS) affects



© The Author(s) 2022. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

nearly 1 in 300 individuals worldwide [1, 2] and is primarily associated with colorectal and endometrial cancers. Heterozygous carriers' risk for these cancers approaches ~80 and ~60%, respectively [3, 4], with onset decades earlier on average compared to sporadic cases [5].

LS is caused by inherited defects in any of four key DNA mismatch repair (MMR) factors: *MSH2*, *MLH1*, *MSH6*, and *PMS2*. These genes are included on most cancer gene panel tests, and it is standard of care to screen them for pathogenic germline variants when their loss is observed in tumors by histology or by tests for microsatellite instability, the mutational consequence of MMR loss. Increased carrier screening in LS holds great potential for reducing risk: it is estimated that the large majority of individuals who carry an LS variant go undetected at present [6], in part because many of them lack the clear family history needed to meet current genetic testing criteria [7, 8].

Despite decades of clinical screening and functional studies of MMR genes, upwards of one-third of the variants identified during clinical genetic evaluation for Lynch syndrome cannot be classified and remain as variants of uncertain significance, or VUS [9, 10]. The eventual reclassification rate for VUS in LS and other hereditary cancer genes is modest, reaching only ~25% [11, 12]. The difficulty of resolving these VUS stands as one of the most persistent barriers to the utility of broader genetic testing for LS genes [13].

As a group, missense variants are particularly challenging to interpret as most are individually rare (i.e., population minor allele frequencies $\leq 10^{-4}$), and they can have functional defects resulting from diverse molecular mechanisms [14]. Over 94% ($n=8614$) of the LS gene missense variants in the NCBI ClinVar database [15] remain as VUS or have conflicting interpretations, reflecting both the volume of VUS' discovery and the challenge of their classification. Even when tumor molecular testing is available, it may not resolve these variants' effects—for example, missense variants can retain protein staining by immunohistochemical testing despite causing mismatch repair deficiency, a source of false negatives that could potentially limit access to immunotherapy [16].

Functionally testing individual missense VUSs in experimental model systems is time- and labor-intensive but can provide key evidence to guide their classification. New approaches (collectively termed multiplex assays of variant effect, or MAVEs) have enabled the systematic testing of many variants at a time [17, 18]. Promisingly, benchmarking comparisons have shown that MAVEs can accurately recapitulate standing classifications from sources such as ClinVar [19, 20]. However, despite the recent proliferation of MAVE studies, practical examples of their use in clinical variant interpretation have been scarce. One recent study [21] evaluated MAVE scores' predictive accuracy for variants found during genetic testing in three cancer-associated genes (*BRCA1*, *TP53*, and *PTEN*). Under recently proposed guidelines [22, 23], the authors were able to reclassify just over half (176/324) of the VUSs that had functional information across the three genes. In another recent study, the same *BRCA1* MAVE scores [24] were intersected with variants discovered during exome sequencing of an unselected healthcare cohort, and an association between MAVE-predicted loss-of-function variants and breast and ovarian cancer diagnoses was observed [25].

Here, we set out to use MAVE-based function maps to facilitate variant classification in Lynch syndrome. We combined a recent MAVE [26] covering 16,749 missense variants in the key Lynch syndrome gene *MSH2* with a clinical dataset containing

15,520 patients with an *LS* gene variant. We validated the *MSH2* MAVE data across 47 previously classified missense variants and found that it meets the established threshold for ‘strong’ functional evidence [27]. Critically, during validation, we excluded any variants for which the standing classification relied upon functional data, thus avoiding the risk of circularity inherent in validating one functional assay using another. We then applied these MAVE scores to 682 standing *MSH2* missense VUS, formally reclassifying 10 to pathogenic/likely pathogenic, and showing that another 497 could be moved to benign/likely benign upon reassessment. Leveraging the detailed, individual-level clinical information in this cohort, we demonstrate that missense *MSH2* variants with abnormal MAVE function scores are associated with elevated colorectal and endometrial cancer risk. Finally, going beyond germline variant interpretation, we demonstrate that MAVE scores can uncover loss-of-function somatic ‘second hits’ as well as biallelic mutations with the availability of tumor DNA tests.

Results

Clinical validation of *MSH2* function map

To establish the clinical validity of multiplexed analyses of variant effect (MAVEs) for interpreting and reclassifying variants in Lynch syndrome, we intersected loss-of-function (LoF) scores from a deep mutational scan of *MSH2* [26] with a clinical database of 1604 individuals with MMR gene variants detected by paired tumor and germline testing. To gauge the strength of evidence provided by our functional data, we curated a list of *MSH2* missense variants previously classified as pathogenic/likely pathogenic ($n=22$) or benign/likely benign ($n=26$) as known controls. Crucially, to avoid the circularity of validating one functional assay using classifications that relied upon other mechanistically similar assays, we included only variants for which clinical interpretation could be reached without the use of any prior functional evidence (e.g., only using population frequency, family history, or tumor characteristics).

Our functional measurements agreed with the clinical interpretation for all 47 of the 47 variants that scored as functionally abnormal or normal (Fig. 1, and Additional file 1: Table S1), with one pathogenic variant scoring in the intermediate range. This resulted in a strength of evidence, as quantified by the OddsPath score [28], of 24.9 for abnormal LoF scores, and 0.043 for neutral scores. Following recommendations for application of the functional evidence criterion using the ACMG/AMP variant interpretation framework [27], *MSH2* LoF scores ≥ 0.4 can therefore be used as

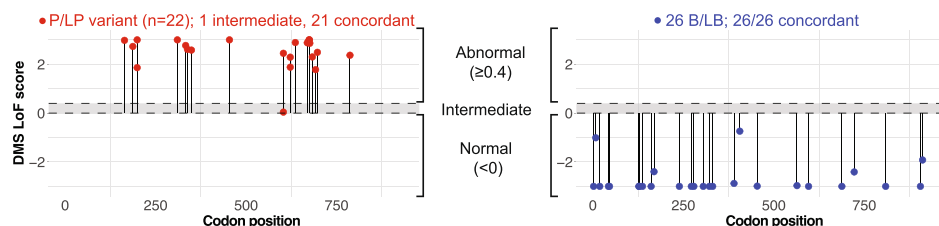


Fig. 1 Validation of *MSH2* function scores across 48 previously classified *MSH2* missense variants. LoF scores for known pathogenic/likely pathogenic (red, at left) and known benign/likely benign variants (blue, at right) are plotted against codon position. Gray shaded interval denotes intermediate score range

‘strong evidence’ in support of variant pathogenicity (PS3 evidence code), while LoF scores < 0 can be conversely be used a ‘strong evidence’ against pathogenicity (BS3 evidence code).

We next applied the *MSH2* missense function scores to a larger panel of individuals ($n=13,916$) for whom germline-only testing had identified at least one germline Lynch syndrome gene variant. Among this cohort, 1937 individuals carried a scorable *MSH2* missense variant. We focused first on the 32 distinct missense variants, carried by 108 individuals, which had been previously classified by the clinical laboratory as pathogenic or likely pathogenic. Of those, 31 had an abnormal function score. The lone exception was the missense variant NM_000251.3(*MSH2*):c.2020G>A (p.Gly674Ser), which was classified as likely pathogenic and has been shown to be partially deficient in ATP binding in vitro [29]; this variant may be a false negative in the functional assay. Of the 31 correctly identified pathogenic variants, 27/31 had an abnormal LoF score from deep mutational scanning, while 4/31 were predicted to be splice disruptive by SpliceAI, indicating that at least among known variants in *MSH2*, disruption at the level of protein function contributes a greater share of the pathogenic burden than splicing defects. In all, the *MSH2* missense function scores achieved a recall of 96.9% in a validation dataset independent from the patient cohort used to derive the OddsPath score.

For the overwhelming majority of *MSH2* missense carriers in this cohort, the variants carried were VUS (1829 individuals, 94.4%). Of the 682 unique such missense VUS, 5.0% scored as functionally abnormal: 24 by DMS LoF score and another 10 by SpliceAI (with another 17 in the intermediate range by either measure; Fig. 2). Thus, loss of function was modestly depleted among these extant variants, relative to the 5130 missense SNVs not observed in this study, among which 6.8% were abnormal (298 and 53 each by DMS and SpliceAI, respectively; $P=0.048$, two-sided binomial test). The depletion of functionally deleterious variants among standing missense VUSs likely reflects the ongoing removal of those with sufficient lines of evidence to be classified as pathogenic.

VUS reclassification

We next pursued clinical variant reclassification for *MSH2* missense variants with abnormal DMS LoF scores, adding functional evidence codes in support of their pathogenicity. Of the 24 such variants, 14 variants had scores exceeding (i.e., more abnormal than) the lowest score in the P/LP validation set (LoF score ≥ 1.7), and we added PS3 (strong evidence) codes for each of these. After adding this evidence, ten of these VUSs met criteria to be reclassified as pathogenic or likely pathogenic (Table 1). Consistent with their causal, pathogenic role, nine of these ten exhibited MMR deficiency as shown by IHC and/or MSI testing. To be conservative, we did not pursue formal reclassification for variants with LoF scores between 0.4 and 1.7; while these were in the abnormal range, they scored below all the training variants used to establish the OddsPath score, and these could in principle be given a weaker evidence code (PS3_moderate) in the future. In sum, there were 14 remaining missense VUSs in the functionally abnormal range which require additional evidence for formal reclassification under the ACMG/AMP framework, including observation in additional cases, or co-occurrence with a somatic loss-of-function variant in the same gene. In the other direction, there were 635 patient missense VUSs (carried by

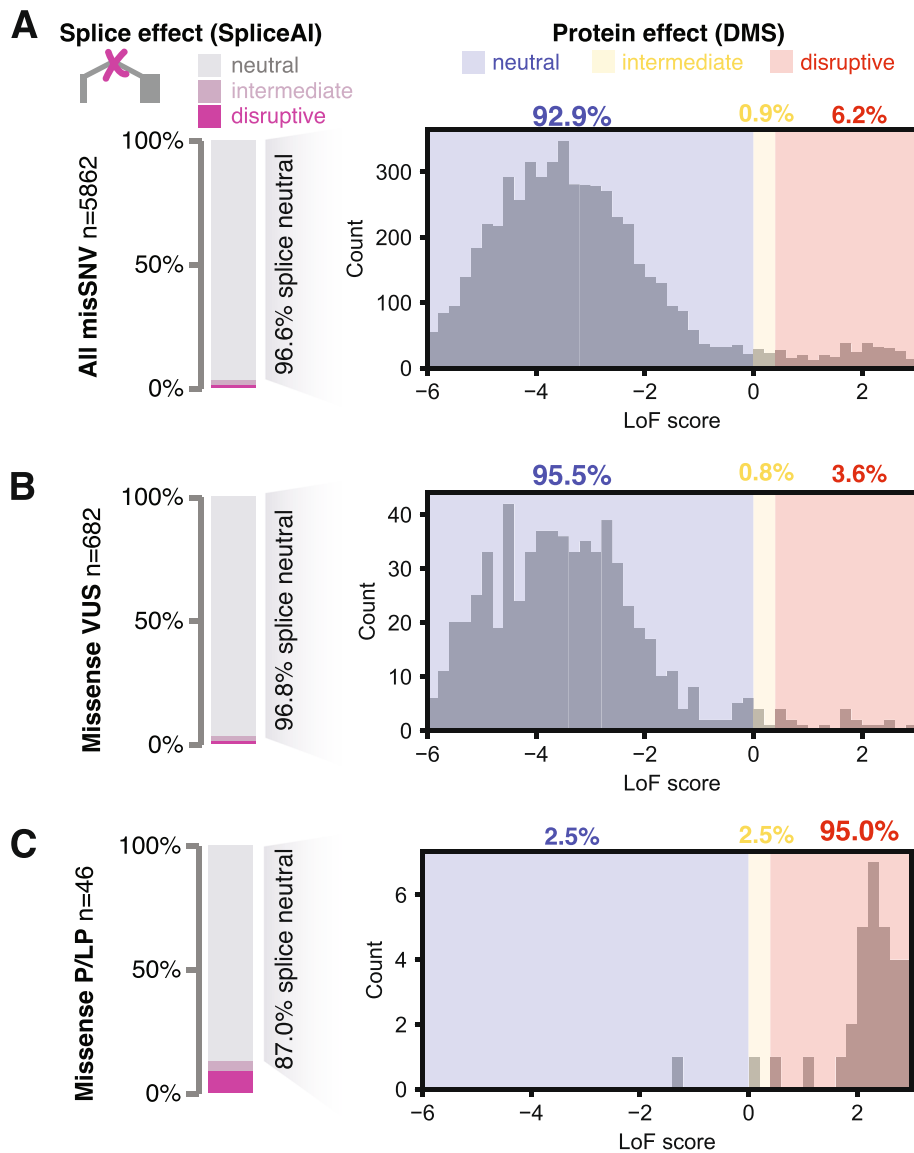


Fig. 2 Function scores and variant reclassification for *MSH2* missense variants, for **A** all single nucleotide missense variants, **B** missense VUSs, and **C** missense variants previously classified as pathogenic or likely pathogenic, including those used for validation. For each group of variants, splicing status was scored by SpliceAI (bar charts at left), and for splice-neutral variants (SpliceAI score<0.2), a histogram of LoF scores from deep mutational scanning are displayed to the right

1772 individuals) which were functionally normal by both deep mutational scanning and SpliceAI prediction. We set out to determine what percentage could potentially be reclassified as benign/likely benign by adding a BS3 functional evidence code and found that 497 of these variants (78.2%) could be reclassified to B/LB with the addition of that evidence. Thus, with the addition of functional evidence, approximately three quarters of all standing *MSH2* missense VUS in this large patient cohort could be newly classified (Fig. 3).

Table 1 Summary of reclassification for missense *MSH2* germline VUS with abnormal LoF scores ($n=24$), with evidence codes applied for each variant

Variant (in NM_000251.3)	Reclassification outcome	LoF Score (Jia et al. 2021) [26]	MAPP MMR (Chao et al. 2008) [30]	gnomAD frequency	LOF score evidence	Pop. database evidence	Clinical (MSI/IHC)	Computational evidence	Additional evidence
c.2083G>A (p.Val695Met)	Reclassified to LP	4.511	20.15	-	PS3	PM2_P	PP4		
c.929T>G (p.Leu310Arg)	Reclassified to LP	3.774	33.02	-	PS3	PM2_P	PP4_S	PP3	PP1
c.1568T>C (p.Phe523Ser)	Reclassified to LP	3.547	10.09	-	PS3	PM2_P	PP4		PM1
c.443T>A (p.Val148Glu)	Insufficient evidence	3.239	4.13	-	PS3	PM2_P	PP4		
c.2011A>T (p.Asn671Tyr)	Reclassified to LP	3.114	20.2	-	PS3	PM2_P	PS4_M	PP3	PM1_P
c.1808A>T (p.Asp603Val)	Reclassified to LP	2.999	22.25	-	PS3	PM2_P	PP4	PP3	
c.1067T>A (p.Ile356Lys)	Reclassified to LP	2.518	14.66	-	PS3	PM2_P	PP4	PP3	
c.2030C>G (p.Thr677Arg)	Reclassified to LP	2.469	37.76	-	PS3	PM2_P	PP4_M	PP3	PM1
c.2276G>T (p.Gly759Val)	Insufficient evidence	2.215	29.36	-	PS3	PM2_P		PP3	
c.2281G>C (p.Gly761Arg)	Reclassified to LP	2.193	40.89	-	PS3	PM2_P			PM1_P
c.2272G>T (p.Asp758Tyr)	Insufficient evidence	1.955	15.52	-	PS3	PM2_P			
c.2099C>A (p.Ala700Glu)	Reclassified to LP	1.917	37.91	-	PS3	PM2_P	PP4		
c.2282G>T (p.Gly761Val)	Insufficient evidence	1.785	9.19	-	PS3	PM2_P	PP4		PM1
c.1241T>C (p.Leu414Pro)	Reclassified to LP	1.766	12.66	-	PS3	PM2_P	BS2_P		
c.2768T>A (p.Val923Glu)	Conflicting/insufficient evidence	1.669	n/a	-	PS3_moderate ^a	PM2_P			
c.1654A>C (p.Thr552Pro)	Insufficient evidence	1.61	10.84	-	PS3_moderate ^a	PM2_P	PS4_M		
c.1601G>C (p.Arg534Pro)	Insufficient evidence	1.383	30.22	-	PS3_moderate ^a	PM2_P			
c.1703C>T (p.Thr568Ile)	Conflicting/insufficient evidence	0.932	4.04	-	PS3_moderate ^a	PM2_P		BP4	
c.1862G>C (p.Arg621Pro)	Insufficient evidence	0.745	30.86	-	PS3_moderate ^a	PM2_P	PP4		
c.2272G>C (p.Asp758His)	Insufficient evidence	0.701	18.79	-	PS3_moderate ^a	PM2_P			
c.386C>G (p.Ser129Cys)	Conflicting/insufficient evidence	0.574	1.84	2.35E-05	PS3_moderate ^a	PM2_P	PP4		BS2_P
c.925G>C (p.Ala309Pro)	Insufficient evidence	0.528	7.68	-	PS3_moderate ^a	PM2_P		PP3	
c.2132G>T (p.Arg711Leu)	Conflicting/insufficient evidence	0.491	19.03	-	PS3_moderate ^a	PM2_P	BS2_P	PP3	
c.1571G>T (p.Arg524Leu)	Insufficient evidence	0.449	17.81	-	PS3_moderate ^a	PM2_P			

^a PS3_moderate codes are applicable for variants with scores between 0.4 and 1.7, but their reclassification was not pursued here

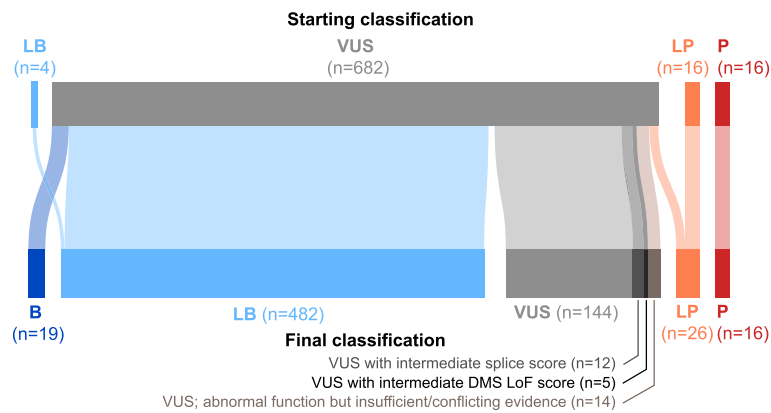


Fig. 3 Reclassification outcomes for 718 *MSH2* missense variants. Flow diagram showing starting and final variant classifications; in total, 74% of the missense VUSs have sufficient evidence to enable potential reclassification to benign (B), likely benign (LB), likely pathogenic (LP) or pathogenic (P). A subset of remaining VUSs had intermediate function scores ($n=19$) or had abnormal function scores but lacked sufficient lines of evidence (or had conflicting evidence) and so remain as VUS ($n=12$)

Cancer prevalence among LOF VUS missense carriers

We next sought to compare the risk conferred by LOF *MSH2* missense variants to that of established P/LP variants in *MSH2* and other LS genes. In this patient cohort, LS-related cancer diagnoses were enriched relative to the general population, but far from completely prevalent: 13.6% of patients in this cohort had a CRC diagnosis, with higher rates in males (38.3%, $n=2229$) compared to females (8.9%, $n=11,687$), possibly reflecting broader inclusion criteria for genetic screening in women (e.g., for breast cancer). Uterine and endometrial cancers (UEC) were similarly prevalent to CRC, affecting 9.5% of females. Other cancers not primarily associated with Lynch syndrome were also prevalent in this cohort, affecting 49.6% of females and 40.7% of males. Gene by gene differences in penetrance closely mirrored those seen previously [2, 31]: P/LP variants in *MLH1* and *MSH2* were the most strongly associated with CRC (odds ratio=14.4 and 8.10, respectively), with lesser effects from *MSH6* and *PMS2* P/LP variants. As previously noted [5], uterine and endometrial cancers differed from colorectal cancer, with *MSH6* (OR=13.2) and *MSH2* (OR=11.9) emerging as the top risk factors, followed by *MLH1* and *PMS2*. Separating *MSH2* missense variants by their functional status, those with abnormal function scores (by DMS or SpliceAI) were significantly associated with both CRC (OR=2.53, 95%CI:[1.04, 6.15], $P=0.04$) and EC (OR=5.56, 95%CI:[2.24,13.8], $P=2.2 \times 10^{-4}$), though with smaller effects than truncating P/LP variants' (Fig. 4). By contrast, *MSH2* missense variants with neutral function scores did not contribute significant risk for CRC or EC ($P \geq 0.67$ for each), nor were they associated with other cancers (Additional file 2: Fig. S1). Thus, loss-of-function missense variants in MMR genes contribute measurable risk for LS-associated cancers, but may exhibit lower risk than their truncating counterparts, underscoring the challenge of their accurate classification.

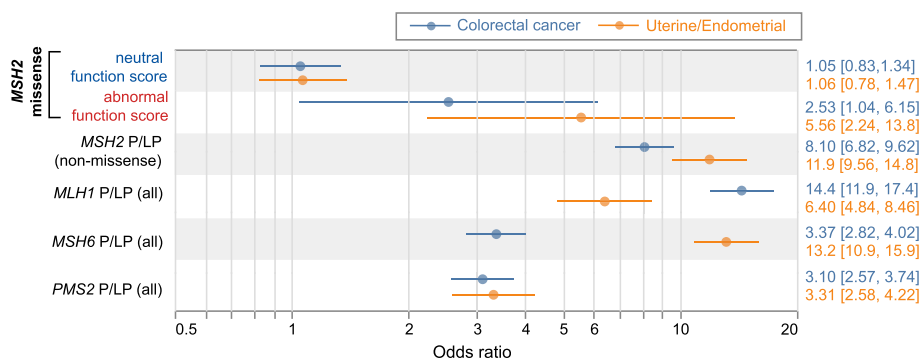


Fig. 4 Cancer associations by variant type. Association between colorectal cancer (blue) or uterine/endometrial cancer (female) and missense variants in *MSH2* (missense, separated by DMS+SpliceAI function score), or P/LP variants in other Lynch syndrome genes; odds ratios shown from logistic regression

Joint annotation of germline and somatic variants

We next sought to apply these functional measures to jointly interpret germline and somatic mutations in *MSH2*. Most cases of Lynch syndrome follow a ‘two hit’ model, with one inherited loss-of-function variant followed by a second somatic mutation disrupting the remaining copy. Therefore, it is expected that in a cohort including individuals with LS, pathogenic somatic ‘second hits’ in *MSH2* would be more common among those individuals who inherited a ‘first hit’ loss-of-function variant in the same gene. We tested this within the paired tumor-normal cohort ($n=1604$ individuals), among the 25 individuals for whom the sole germline finding was a missense *MSH2* variant (Fig. 5 and Additional file 3: Table S2). DMS scores indicated 13 of these 25 germline variants are functionally deleterious, constituting pathogenic inherited ‘first hits’. Among these 13 carriers, 12 (92.4%) had a P/LP somatic ‘second hit’ in *MSH2*, or a structural variant in the upstream gene *EPCAM* (which causes epigenetic silencing of *MSH2*) [32, 33]. By contrast, among the other 12 individuals who inherited a single *MSH2* missense variant scored as neutral by DMS, only two (16.7%) had a P/LP somatic mutation in *MSH2*, a significantly lower prevalence ($P=0.00021$, Fisher’s exact test).

We next examined these functional measures’ association with tumor microsatellite instability (MSI), a hallmark of MMR deficiency. By genotyping at microsatellite markers, tumors can be rated as microsatellite stable (MSS), MSI-low, or MSI-high, with distinct implications for treatment and prognosis [34]. We observed that MSI was universal among patients with *MSH2* missense alleles that had abnormal function scores in our map, reflecting their functional disruptiveness. After excluding individuals with *MLH1* promoter hypermethylation (an independent somatic epigenetic mechanism sufficient to cause MMR deficiency and MSI), tumors from all 11 of 11 individuals who carried a germline *MSH2* missense variant deemed LoF by DMS were MSI-high, whereas only 2/6 individuals with a germline functionally neutral missense *MSH2* variant and without *MLH1* hypermethylation had MSI-high or MSI-low tumors ($P=0.006$), and both of those two cases could be explained by biallelic somatic mutations in other LS genes (in one, NM_000535.7(*PMS2*):c.1687C>T (p.Arg563Ter) + LOH; in the other, NM_000249.4(*MLH1*):c.676C>T (p.Arg226Ter) + LOH). The ability of the *MSH2* function map to identify germline variants associated

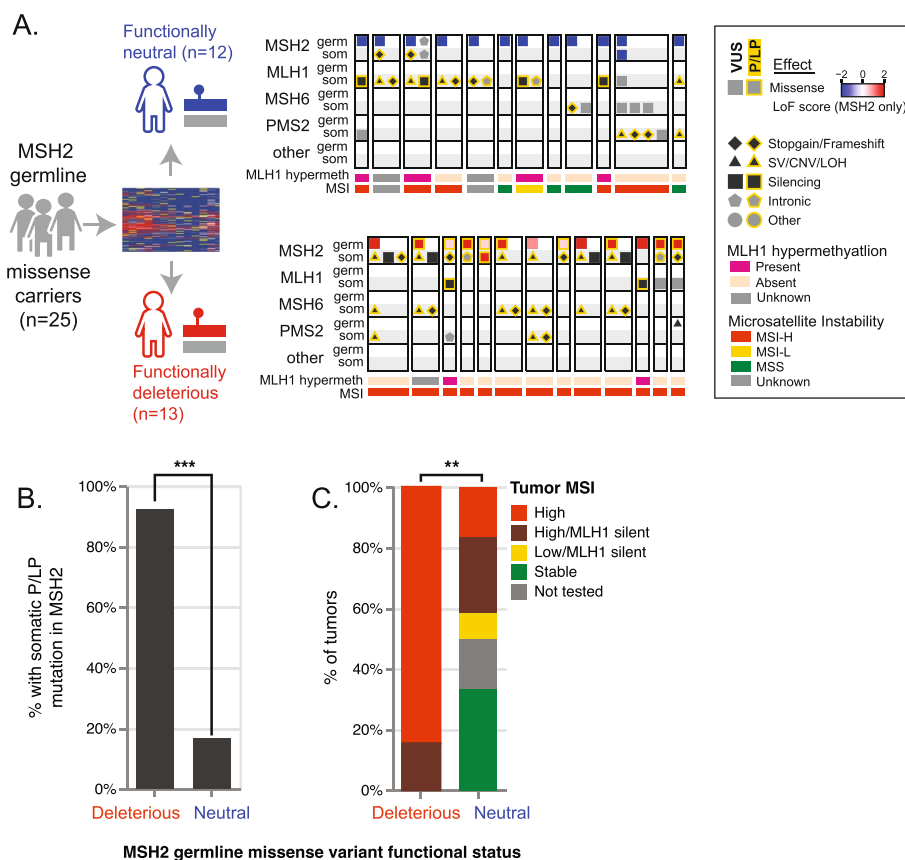


Fig. 5 Joint analysis of germline and tumor mutations. **A** Patterns of germline and somatic mutations among LS genes, among germline carriers of *MSH2* missense variants, separated into those scored as functionally neutral (top) or deleterious (bottom) by deep mutational scanning LoF score. **B** Fraction of individuals with a somatic P/LP mutation in *MSH2*, by *MSH2* germline missense functional status. **C** Tumor microsatellite status, by *MSH2* germline missense variant functional status. ***, $P < 0.001$; **, $P < 0.01$

with pathogenic somatic second hits, and MSI demonstrates how patterns of tumor mutation can support germline variant classification [35].

Somatic mutations also contribute to the variant interpretation burden, particularly for heavily mutated MMR-deficient tumors. Indeed, among the 437 individuals for which paired testing revealed at least one somatic *MSH2* mutation, most (382, 87.4%) had more than one somatic mutation in a tested gene, and nearly half (182, 47.6%) had multiple somatic mutations in *MSH2* alone. We focused on the 84 individuals who carried at least one missense somatic *MSH2* variant, comparing the 46 carrying at least one somatic missense *MSH2* variant we predict to be functionally disruptive, and the 38 for whom these somatic missense variants were exclusively functionally neutral and who did not carry any other somatic P/LP *MSH2* mutations (Fig. 6 and Additional file 4: Table S3). Notably, among the latter (carriers of exclusively functionally neutral somatic *MSH2* variants), additional somatic mutations in other LS genes (i.e., *MSH6*, *MLH1*, *PMS2*) were found in all 38 tumors (100%). By contrast, when at least one functionally disruptive *MSH2* missense somatic mutation was found, somatic mutations in other LS genes were significantly less common (30 of 46 tumors; Fisher’s exact $P = 2.70 \times 10^{-5}$). Similarly, *MLH1* promoter hypermethylation was present in nearly a third of the tumors

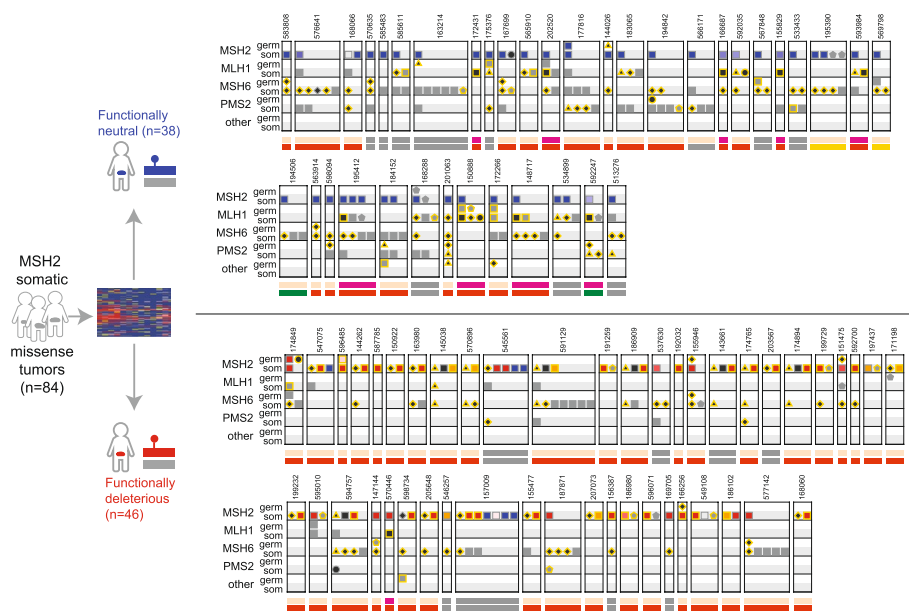


Fig. 6 Mutational patterns in patients with somatic *MSH2* missense variants. Tumor and germline mutations in LS genes are shown in patients who carry functionally neutral somatic *MSH2* missense variants (upper track, $n=38$), or those with a functionally abnormal somatic *MSH2* variant (lower track, $n=46$). Mutations and tumor characteristics are denoted as in Fig. 5A

for which the only somatic *MSH2* mutations were functionally neutral missense (9 of the 28 tumors in which *MLH1* was assayed, 32.1%) but nearly absent among tumors with at least one somatic missense *MSH2* mutation deemed LoF by DMS/SpliceAI (1 of 38, 2.6%; Fisher’s exact $P=0.0013$). MAVE measurements can thus identify functionally disruptive somatic mutations driving tumor MMR deficiency even in the absence of an inherited loss-of-function variant.

Discussion

As MAVE function maps are put into practice for clinical variant interpretation, an important prerequisite is to assess their predictive value for disease risk and clinical phenotypes. Here we did so in the context of *MSH2*, a key DNA repair factor underlying Lynch syndrome. We supplemented protein-level MAVE effect measurements with deep learning-based splicing effect predictions [36], to newly reclassify over 74% of the 682 missense variants of uncertain significance (VUS) encountered in *MSH2* in a cohort of tens of thousands of individuals with genetic testing results.

In particular, the reclassification of 10 variants to pathogenic/likely pathogenic newly enabled the return of definitive genetic diagnoses. Reclassification of these variants as pathogenic has critical clinical implications for these patients and family members who also inherited them, such as more frequent colonoscopies, risk-reducing surgeries to avoid gynecologic cancers, initiation of esophagogastroduodenoscopy for upper GI cancer surveillance and additional cancer screening recommendations not necessarily performed for the general population. Going forward, these MAVE-based function scores are now integrated into the variant interpretation process at clinical genetic testing

laboratories and will assist with classification of newly observed rare *MSH2* missense variants.

Our study leverages several unique features of this large cohort. Firstly, to validate the *MSH2* MAVE, we selected a set of control variants for which the classification stands without including prior functional evidence, that is, based upon orthogonal features such as recurrence, tumor characteristics, and co-segregation with early-onset cancer. This avoids the risk of validating a MAVE in part by prior functional evidence from existing assays, which despite being lower-throughput, may be mechanistically similar and highly correlated with the MAVE.

For many genes, culling the training set to remove such variants may not be practical—obtaining a sufficient number of control variants is emerging as a key rate limiting step for many MAVEs; at least 11 control variants are needed to reach a ‘moderate’ strength of evidence [27]. This challenge is highlighted by a recent effort to reclassify variants in *PTEN* using MAVE data [21], which was hampered by the limited number ($n=2$) of known benign variants. In many cases, filtering variants for this or other criteria may not even be possible: public-facing databases such as ClinVar are often the primary source for these controls, and to protect privacy, they do not provide individual-level clinical or demographic data.

We used per-individual clinical information to explore the association between loss of function, as indicated by MAVE scores, and cancer prevalence. We observed that *MSH2* missense variants with abnormal function identified by MAVE were associated with significantly elevated risk for LS-associated colorectal and uterine/endometrial cancers. Notably, these associations were weaker than those observed for P/LP variants that were not missense (i.e., truncating frameshift or stop-gain variants). Thus, functionally abnormal *MSH2* missense variants as a group may be less penetrant than their truncating counterparts, while still being measurably pathogenic within the population of individuals selected for germline cancer testing.

In contrast, another recent study of MMR gene variant carriers found no difference in the incidence of LS-related cancers between carriers of *MSH2* missense P/LP and truncating variants [37]. A key difference, however, was that the missense variants included in that study were restricted to those with standing P/LP classifications, which may reflect a particularly severe subset. Thus, the addition of MAVE-based functional data may have captured missense variants with intermediate functional defects which confer a moderate level of risk. An important future direction will be to replicate this analysis in an unselected population, as has recently been done for *BRCA1* [25], and to model polygenic risk as a potential modifier [38].

To date, applications of MAVE data have largely focused on germline variants. Here, we demonstrated how MAVEs can also support joint analyses of germline and somatic mutations, leveraging a clinical database of 1604 individuals with paired tumor-normal tests. As expected under Knudson’s two-hit hypothesis [39], among individuals who inherited an *MSH2* missense variant, pathogenic somatic ‘second hits’ in *MSH2* were significantly more common when the MAVE data indicated the germline variant was functionally disruptive as compared to normal. In addition, we identified 29 individuals whose cancers had double *MSH2* somatic mutations with at least one of the mutations identified as functionally deleterious

by MAVE. Excluding an inherited predisposition as the cause for these individuals' MMRd tumors has the potential to prevent unnecessary screenings for their blood relatives.

The large majority of *MSH2* missense VUS considered here are functionally normal and do not provide a basis for a positive diagnosis. Nevertheless, in the context of an affected individual, they may still be partially informative by suggesting that 'causal' variant(s) may reside at a different locus. Alternatively, they may suggest a different molecular mechanism, including somatic mutation, epigenetic silencing (e.g., *MLH1* promoter hypermethylation or loss of *MSH2* expression secondary to *EPCAM* 3' deletions), or structural variants [40, 41], some of which are challenging to detect by standard pipelines. Indeed, we observed that germline carriers of functionally normal *MSH2* missense variants had much higher rates of *MLH1* disruption by silencing and/or somatic mutation (Fig. 5) relative to carriers of disruptive *MSH2* missense variants. Likewise, individuals with functionally normal somatic missense mutations in *MSH2* had a much higher rate of somatic disruption of the other three primary LS factors (Fig. 6). As MAVE function maps become more broadly available, they may allow previously suspected VUS to be ruled out, offering an opportunity to identify previously obscured functional variants elsewhere.

A limitation of this study is that the MAVE function scores used here were derived from a cDNA-based deep mutational scan and so do not capture splice disruptive effects, which although in the minority relative to protein-disruptive variants, may still account for a substantial number of cases for LS [42, 43]. These effects can be obtained experimentally with other MAVE approaches such as saturation genome editing [24] or saturation prime editing [44], or directly measured by massively parallel splicing assays [45–47]. For the purposes of this study, we used predictions from SpliceAI [36], a deep learning-based splicing effect predictor which has been shown to be highly accurate [48].

Conclusions

As gene panel and exome sequencing are increasingly utilized in the clinical setting for a variety of indications, there is an opportunity to leverage the massive scale of MAVE experiments to prospectively generate functional evidence for as yet unseen variants. Here we established the validity of MAVE-based functional evidence for missense variant classification in the Lynch syndrome gene *MSH2*, and demonstrated how these functional measures can support resolution of standing missense VUS. With proper clinical validation, it appears promising that MAVE data may soon play a primary role in identifying patients who may not have otherwise come to clinical attention, but could benefit from additional monitoring based upon their genetic risk.

Methods

Patient population

Clinical information and genetic variants were obtained for patients found to carry at least one variant in any of the four major Lynch syndrome genes (*MSH2*, *MLH1*, *MSH6*,

PMS2) during multi-gene panel testing for cancer predisposition at Ambry Genetics before December 14, 2020. We obtained data for 13,916 LS gene variant carriers who underwent germline-only testing, and another 1604 patients who underwent paired tumor-germline testing at Ambry Genetics before August 31, 2020.

Functional annotation of *MSH2* missense variants

Each *MSH2* missense variant was annotated with two function scores: the loss-of-function (LoF) scores from a recent deep mutational scan [26] which measures impact upon *MSH2* protein function, and SpliceAI deltaMax scores [36], a computational estimate of the probability of splicing disruption. Variants with an LoF score ≥ 0.4 or a SpliceAI deltaMax score ≥ 0.5 were considered deleterious; those with LoF scores between 0 and 0.4 or deltaMax between 0.2 and 0.5 were considered intermediate, while those with LoF scores < 0 and deltaMax < 0.2 were considered functionally neutral.

Variant classification

Patient variant classification was performed at Ambry Genetics using a point-based implementation [49] of ACMG/AMP variant classification guidelines [23, 50], assigning each variant into one of five tiers: pathogenic (P), likely pathogenic (LP), uncertain significance (VUS), likely benign (LB) or benign (B). To validate the *MSH2* missense function scores, we used previously classified *MSH2* missense variants from the paired (tumor-normal) dataset. To avoid redundant application of evidence, we used only those variants which had sufficient evidence to be classified as benign/likely benign and pathogenic/likely pathogenic without use of any prior functional data. Function scores' strength of evidence for or against pathogenicity was quantified using the Oddspath score [28]. Additionally, structural evidence was assessed using a standard structural modelling protocol and energies of destabilization compared to nearby informative variants and identification of impacted motifs [51, 52].

Cancer association

For analysis of cancer prevalence among LS variant carriers, patients were categorized by variant classification(s) and gene(s) affected. For analysis of *MSH2* VUS carriers, we excluded individuals who also carried a pathogenic or likely pathogenic (P/LP) variant in a non-Lynch syndrome gene, while individuals with both an *MSH2* VUS and a P/LP variant in *MLH1*, *MSH6*, or *PMS2* were considered as carriers for those respective genes and were excluded from *MSH2* VUS association tests. Logistic regression models were fit using the python statsmodels package version 0.12.2 [53] using cancer diagnosis as the response variable and, as features, each individual's carrier status for the following categories of variants, each encoded as zero or one: (1) *MSH2* missense with deleterious function score, (2) *MSH2* missense with neutral function score, (3) *MSH2* other P/LP, (4) *MLH1* any P/LP, (5) *MSH6* any P/LP, and (6) *PMS2* any P/LP. Models were fit separately for colorectal cancer and uterine/endometrial cancer (in the latter case, including only females).

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13059-022-02839-z>.

Additional file 1: Table S1. Validation variants used to calculate OddsPath score.

Additional file 2: Figure S1. Associations with other cancers, by variant type.

Additional file 3: Table S2. Variants identified by paired tumor-normal test, among patients with a single germline *MSH2* variant.

Additional file 4: Table S3. Tumor and germline mutations in individuals with *MSH2* somatic tumor mutations.

Additional file 5. Review history.

Acknowledgements

We thank Matthew Varga and Min-Sun Park of Ambry Genetics for support of the structural investigation of these variants, and members of the Kitzman lab for helpful comments.

Review history

The review history is available as Additional file 5.

Peer review information

Anahita Bishop was the primary editor of this article and managed its editorial process and peer review in collaboration with the rest of the editorial team.

Authors' contributions

F.H. and R.K. abstracted clinical information. A.S., F.H., A.C., C.S., and J.O.K. analyzed the data. A.S., F.H., R.K., and J.O.K. wrote the manuscript. All author(s) read and approved the final manuscript.

Authors' information

Twitter handles: @rachidkaram (Rachid Karam), @jacobkitzman (Jacob O. Kitzman)

Funding

This work was supported by the National Institute of General Medical Sciences (R01GM129123 to J.O.K.).

Availability of data and materials

MSH2 loss-of-function scores are available at MaveDB (<https://www.mavedb.org>) under accession urn:mavedb:00000050-a and are also provided as Table S4 in the published *MSH2* deep mutational scan [26]. Underlying sequencing counts are available from NCBI GEO under accession GSE162130 [26]. MMR gene variants and clinical interpretation are deposited in NCBI ClinVar: [https://www.ncbi.nlm.nih.gov/clinvar/?gr=0&term=\(\(msh2%5BGene%20Name%5D\)%20OR%20msh6%5BGene%20Name%5D\)%20OR%20mlh1%5BGene%20Name%5D\)%20OR%20pms2%5BGene%20Name%5D](https://www.ncbi.nlm.nih.gov/clinvar/?gr=0&term=((msh2%5BGene%20Name%5D)%20OR%20msh6%5BGene%20Name%5D)%20OR%20mlh1%5BGene%20Name%5D)%20OR%20pms2%5BGene%20Name%5D) SpliceAI software: <https://github.com/Illumina/SpliceAI> [36].

Declarations

Ethics approval and consent to participate

Data collection and sharing procedures were reviewed by the Western IRB and University of Michigan Medical School IRB (study HUM00220511) and declared exempt from human subject regulations. All experimental methods comply with the Helsinki Declaration.

Consent for publication

Not applicable.

Competing interests

F.H., A.C., and R.K. are employees of Ambry Genetics. J.O.K. serves as a scientific advisor to MyOme, Inc. The authors declare that there are no further competing interests.

Received: 15 August 2022 Accepted: 13 December 2022

Published online: 22 December 2022

References

1. Win AK, Jenkins MA, Dowty JG, Antoniou AC, Lee A, Giles GG, et al. Prevalence and penetrance of major genes and polygenes for colorectal cancer. *Cancer Epidemiol Biomark Prev.* 2017;26:404–12.
2. Haraldsdottir S, Rafnar T, Frankel WL, Einarsdottir S, Sigurdsson A, Hampel H, et al. Comprehensive population-wide analysis of Lynch syndrome in Iceland reveals founder mutations in *MSH6* and *PMS2*. *Nat Commun.* 2017;8:14755.
3. Jaspersion KW, Tuohy TM, Neklason DW, Burt RW. Hereditary and familial colon cancer. *Gastroenterology.* 2010;138:2044–58.
4. Moreira L, Balaguer F, Lindor N, de la Chapelle A, Hampel H, Aaltonen LA, et al. Identification of Lynch syndrome among patients with colorectal cancer. *JAMA.* 2012;308:1555–65.

5. Dominguez-Valentin M, Sampson JR, Seppälä TT, ten Broeke SW, Plazzer J-P, Nakken S, et al. Cancer risks by gene, age, and gender in 6350 carriers of pathogenic mismatch repair variants: findings from the Prospective Lynch Syndrome Database. *Genet Med*. 2019;22:15–25.
6. Hampel H, de la Chapelle A. The search for unaffected individuals with Lynch syndrome: do the ends justify the means? *Cancer Prev Res*. 2011;4:1–5.
7. Sjursen W, Haukanes BI, Grindedal EM, Aarset H, Stormorken A, Engebretsen LF, et al. Current clinical criteria for Lynch syndrome are not sensitive enough to identify MSH6 mutation carriers. *J Med Genet*. 2010;47:579–85.
8. LaDuca H, Polley EC, Yussuf A, Hoang L, Gutierrez S, Hart SN, et al. A clinical guide to hereditary cancer panel testing: evaluation of gene-specific cancer associations and sensitivity of genetic testing criteria in a cohort of 165,000 high-risk patients. *Genet Med*. 2020;22:407–15.
9. Sijmons RH, Greenblatt MS, Genuardi M. Gene variants of unknown clinical significance in Lynch syndrome. An introduction for clinicians. *Familial Cancer*. 2013;12:181–7.
10. Tricarico R, Kasela M, Mareni C, Thompson BA, Drouet A, Staderini L, et al. Assessment of the InSiGHT Interpretation Criteria for the Clinical Classification of 24 MLH1 and MSH2 Gene Variants. *Hum Mutat*. 2017;38:64–77.
11. Mersch J, Brown N, Pirzadeh-Miller S, Mundt E, Cox HC, Brown K, et al. Prevalence of variant reclassification following hereditary cancer genetic testing. *JAMA*. 2018;320:1266–74.
12. Welsh JL, Hoskin TL, Day CN, Thomas AS, Cogswell JA, Couch FJ, et al. Clinical decision-making in patients with variant of uncertain significance in BRCA1 or BRCA2 Genes. *Ann Surg Oncol*. 2017;24:3067–72.
13. Hampel H, Yurgelun MB. Point/counterpoint: is it time for universal germline genetic testing for all GI cancers? *J Clin Oncol*. 2022;JCO2102764.
14. Backwell L, Marsh JA. Diverse molecular mechanisms underlying pathogenic protein mutations: beyond the loss-of-function paradigm. *Annu Rev Genomics Hum Genet*. 2022. <https://doi.org/10.1146/annurev-genom-111221-103208>.
15. Landrum MJ, Lee JM, Benson M, Brown GR, Chao C, Chitipiralla S, et al. ClinVar: improving access to variant interpretations and supporting evidence. *Nucleic Acids Res*. 2018;46:D1062–7.
16. Hechtman JF, Rana S, Middha S, Stadler ZK, Latham A, Benayed R, et al. Retained mismatch repair protein expression occurs in approximately 6% of microsatellite instability-high cancers and is associated with missense mutations in mismatch repair genes. *Mod Pathol*. 2020;33:871–9.
17. Weile J, Roth FP. Multiplexed assays of variant effects contribute to a growing genotype-phenotype atlas. *Hum Genet*. 2018;137:665–78.
18. Starita LM, Ahituv N, Dunham MJ, Kitzman JO, Roth FP, Seelig G, et al. Variant interpretation: functional assays to the rescue. *Am J Hum Genet*. 2017;101:315–25.
19. Livesey BJ, Marsh JA. Using deep mutational scanning to benchmark variant effect predictors and identify disease mutations. *Mol Syst Biol*. 2020;16:e9380.
20. Cubuk C, Garrett A, Choi S, King L, Loveday C, Torr B, et al. Clinical likelihood ratios and balanced accuracy for 44 in silico tools against multiple large-scale functional assays of cancer susceptibility genes. *Genet Med*. 2021. <https://doi.org/10.1038/s41436-021-01265-z>.
21. Fayer S, Horton C, Dines JN, Rubin AF, Richardson ME, McGoldrick K, et al. Closing the gap: Systematic integration of multiplexed functional data resolves variants of uncertain significance in BRCA1, TP53, and PTEN. *Am J Hum Genet*. 2021;108:2248–58.
22. Gelman H, Dines JN, Berg J, Berger AH, Brnich S, Hisama FM, et al. Recommendations for the collection and use of multiplexed functional data for clinical variant interpretation. *Genome Med*. 2019;11:85.
23. Brnich SE, Rivera-Muñoz EA, Berg JS. Quantifying the potential of functional evidence to reclassify variants of uncertain significance in the categorical and Bayesian interpretation frameworks. *Hum Mutat*. 2018;39:1531–41.
24. Findlay GM, Daza RM, Martin B, Zhang MD, Leith AP, Gasperini M, et al. Accurate classification of BRCA1 variants with saturation genome editing. *Nature*. 2018;562:217–22.
25. Schiabor Barrett KM, Masnick M, Hatchell KE, Savatt JM, Banet N, Buchanan A, et al. Clinical validation of genomic functional screen data: analysis of observed BRCA1 variants in an unselected population cohort. *HGG Adv*. 2022;3:100086.
26. Jia X, Burugula BB, Chen V, Lemons RM, Jayakody S, Maksutova M, et al. Massively parallel functional testing of MSH2 missense variants conferring Lynch syndrome risk. *Am J Hum Genet*. 2021;108:163–75.
27. Brnich SE, Abou Tayoun AN, Couch FJ, Cutting GR, Greenblatt MS, Heinen CD, et al. Recommendations for application of the functional evidence PS3/BS3 criterion using the ACMG/AMP sequence variant interpretation framework. *Genome Med*. 2019;12:3.
28. Tavtigian SV, Greenblatt MS, Harrison SM, Nussbaum RL, Prabhu SA, Boucher KM, et al. Modeling the ACMG/AMP variant classification guidelines as a Bayesian classification framework. *Genet Med*. 2018;20:1054–60.
29. Heinen CD, Wilson T, Mazurek A, Berardini M, Butz C, Fishel R. HNPCC mutations in hMSH2 result in reduced hMSH2-hMSH6 molecular switch functions. *Cancer Cell*. 2002;1:469–78.
30. Chao EC, Velasquez JL, Witherspoon MS, Rozek LS, Peel D, Ng P, Gruber SB, Watson P, Rennett G, Anton-Culver H, Lynch H, Lipkin SM. Accurate classification of MLH1/MSH2 missense variants with multivariate analysis of protein polymorphisms-mismatch repair (MAPP-MMR). *Hum Mutat*. 2008;29(6):852–60. <https://doi.org/10.1002/humu.20735>.
31. Møller P, Seppälä TT, Bernstein I, Holinski-Feder E, Sala P, Gareth Evans D, et al. Cancer risk and survival in path_MMR carriers by gene and gender up to 75 years of age: a report from the Prospective Lynch Syndrome Database. *Gut*. 2018;67:1306–16.
32. Niessen RC, Hofstra RMW, Westers H, Ligtenberg MJL, Kooi K, Jager POJ, et al. Germline hypermethylation of MLH1 and EPCAM deletions are a frequent cause of Lynch syndrome. *Genes Chromosom Cancer*. 2009;48:737–44.
33. Ligtenberg MJL, Kuiper RP, Chan TL, Goossens M, Hebeda KM, Voorendt M, et al. Heritable somatic methylation and inactivation of MSH2 in families with Lynch syndrome due to deletion of the 3' exons of TACSTD1. *Nat Genet*. 2009;41:112–7.

34. Boland CR, Thibodeau SN, Hamilton SR, Sidransky D, Eshleman JR, Burt RW, et al. A National Cancer Institute Workshop on Microsatellite Instability for cancer detection and familial predisposition: development of international criteria for the determination of microsatellite instability in colorectal cancer. *Cancer Res.* 1998;58:5248–57.
35. Shirts BH, Konnick EQ, Upham S, Walsh T, Ranola JMO, Jacobson AL, et al. Using somatic mutations from tumors to classify variants in mismatch repair genes. *Am J Hum Genet.* 2018;103:19–29.
36. Jaganathan K, Kyriazopoulou Panagiotopoulou S, McRae JF, Darbandi SF, Knowles D, Li YI, et al. Predicting splicing from primary sequence with deep learning. *Cell.* 2019;176:535–548.e24.
37. Dominguez-Valentin M, Plazzer J-P, Sampson JR, Engel C, Aretz S, Jenkins MA, et al. No difference in penetrance between truncating and missense/aberrant splicing pathogenic variants in MLH1 and MSH2: a Prospective Lynch Syndrome Database Study. *J Clin Med Res.* 2021;10. <https://doi.org/10.3390/jcm10132856>.
38. Fahed AC, Wang M, Homburger JR, Patel AP, Bick AG, Neben CL, et al. Polygenic background modifies penetrance of monogenic variants for tier 1 genomic conditions. *Nat Commun.* 2020;11:3635.
39. Knudson AG Jr. Mutation and cancer: statistical study of retinoblastoma. *Proc Natl Acad Sci U S A.* 1971;68:820–3.
40. Rhees J, Arnold M, Boland CR. Inversion of exons 1–7 of the MSH2 gene is a frequent cause of unexplained Lynch syndrome in one local population. *Familial Cancer.* 2014;13:219–25.
41. Frytchard CC, Morrissey C, Kumar A, Zhang X, Smith C, Coleman I, et al. Complex MSH2 and MSH6 mutations in hypermutated microsatellite unstable advanced prostate cancer. *Nat Commun.* 2014;5:4988.
42. Rhine CL, Cygan KJ, Soemedi R, Maguire S, Murray MF, Monaghan SF, et al. Hereditary cancer genes are highly susceptible to splicing mutations. *PLoS Genet.* 2018;14:e1007231.
43. Morak M, Pineda M, Martins A, Gaildrat P, Tubeuf H, Drouet A, et al. Splicing analyses for variants in MMR genes: best practice recommendations from the European Mismatch Repair Working Group. *Eur J Hum Genet.* 2022;1–9.
44. Erwood S, Bily TMI, Lequyer J, Yan J, Gulati N, Brewer RA, et al. Saturation variant interpretation using CRISPR prime editing. *Nat Biotechnol.* 2022;40:885–95.
45. Gergics P, Smith C, Bando H, Jorge AAL, Rockstroh-Lippold D, Vishnopolska SA, et al. High-throughput splicing assays identify missense and silent splice-disruptive POU1F1 variants underlying pituitary hormone deficiency. *Am J Hum Genet.* 2021;108:1526–39.
46. Adamson SI, Zhan L, Graveley BR. Vex-seq: high-throughput identification of the impact of genetic variation on pre-mRNA splicing efficiency. *Genome Biol.* 2018;19:71.
47. Soemedi R, Cygan KJ, Rhine CL, Wang J, Bulacan C, Yang J, et al. Pathogenic variants that alter protein code often disrupt splicing. *Nat Genet.* 2017;49:848–55.
48. Rentzsch P, Schubach M, Shendure J, Kircher M. CADD-Splice-improving genome-wide variant effect prediction using deep learning-derived splice scores. *Genome Med.* 2021;13:31.
49. Pesaran T, Karam R, Huether R, Li S, Farber-Katz S, Chamberlin A, et al. Beyond DNA: an integrated and functional approach for classifying germline variants in breast cancer genes. *Int J Breast Cancer.* 2016;2016:2469523.
50. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med.* 2015;17:405–24.
51. Martin S, Chamberlin A, Shinde DN, Hempel M, Strom TM, Schreiber A, et al. De novo variants in GRIA4 lead to intellectual disability with or without seizures and gait abnormalities. *Am J Hum Genet.* 2017;101:1013–20.
52. Sherrill JD, Kc K, Wang X, Wen T, Chamberlin A, Stucke EM, et al. Whole-exome sequencing uncovers oxidoreductases DHTKD1 and OGDHL as linkers between mitochondrial dysfunction and eosinophilic esophagitis. *JCI Insight.* 2018;3. <https://doi.org/10.1172/jci.insight.99922>.
53. Seabold S, Perktold J. Statsmodels: Econometric and statistical modeling with python. Proceedings of the 9th Python in Science Conference [Internet]. SciPy; 2010. Available from: <https://conference.scipy.org/proceedings/scipy2010/seabold.html>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

