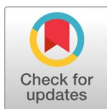


BioScientific Review (BSR)

Volume 8 Issue 3, 2026

ISSN(P): 2663-4198, ISSN(E): 2663-4201

Homepage: <https://journals.umt.edu.pk/index.php/bsr>



- Title:** **In silico Analysis of Non-Synonymous Single Nucleotide Polymorphisms (nsSNPs) in the SLC40A1 Gene Associated with Hemochromatosis Type 4**
- Author (s):** Rimsha Zafar¹, Naveera Khan², Muhammad Awais¹, and Zartash Butt³
- Affiliation (s):** ¹University of Central Punjab, Lahore, Pakistan
²Quaid-e-Azam University, Islamabad, Pakistan
³University of the Punjab, Lahore, Pakistan
- DOI:** <https://doi.org/10.32350/bsr.83.03>.
- History:** Received: May 09, 2026, Revised: July 20, 2026, Accepted: August 10, 2026, Published: September 21, 2026
- Citation:** Zafar R, Khan N, Awais M, Butt Z. In silico analysis of non-synonymous single nucleotide polymorphisms (nsSNPs) in the SLC40A1 gene associated with hemochromatosis type 4. *BioSci Rev.* 2026;8(3):42-56.
<https://doi.org/10.32350/bsr.83.03>.
- Copyright:** © The Authors
- Licensing:**  This article is open access and is distributed under the terms of [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/)
- Conflict of Interest:** Author(s) declared no conflict of interest



UMT

A publication of
The Department of Life Sciences, School of Science
University of Management and Technology, Lahore, Pakistan

***In silico* Analysis of Non-Synonymous Single Nucleotide Polymorphisms (nsSNPs) in the *SLC40A1* Gene Associated with Hemochromatosis Type 4**

Rimsha Zafar^{1*}, Naveera Khan², Muhammad Awais¹, and Zartash Butt³

¹Department of Biotechnology, Faculty of Science and Technology, University of Central Punjab, Lahore, Pakistan

²Department of Biotechnology, Quaid-e-Azam University, Islamabad, Pakistan

³Centre for Applied Molecular Biology, University of Punjab, Lahore, Pakistan

ABSTRACT

Background. Variants in *SLC40A1*, which encodes ferroportin (FPN), the only known mammalian cellular iron exporter, cause hemochromatosis type 4 (HH4), or ferroportin disease, a rare autosomal dominant disorder of iron metabolism. Non-synonymous single nucleotide polymorphisms (nsSNPs) in *SLC40A1* may impair FPN folding, membrane localization, iron efflux or hepcidin responsiveness. Yet, researchers have not functionally annotated most catalogued variants. This *in silico* study systematically screened all coding-region nsSNPs of *SLC40A1* using an integrated bioinformatics pipeline to priorities high-confidence deleterious candidates for experimental validation.

Methodology. The study retrieved 300 coding-region nsSNPs (RefSeq NM_014585.6; UniProt Q9NP59; dbSNP build 156) from NCBI dbSNP and Ensembl. Furthermore, the study assessed functional impact with PolyPhen-2, PROVEAN, SIFT and PhD-SNP; thermodynamic stability with I-Mutant 3.0 and MUpro; residue-level effects with Project HOPE; and secondary/tertiary structure with PSIPRED, I-TASSER and GalaxyRefine. We retained variants that all four functional predictors called deleterious and that at least one stability tool called destabilising.

Results. PolyPhen-2 classified 32.6% of nsSNPs as probably damaging, PROVEAN 33.6% as deleterious, SIFT 31.3% as not tolerated, and PhD-SNP 22.0% as disease-associated. I-Mutant 3.0 and MUpro predicted destabilisation for 32.3% and 33.7% of variants, respectively. Consensus filtering yielded 55 high-confidence deleterious nsSNPs; HOPE indicated that 38 substantially perturb FPN structure/function and 17 produced moderate effects. A total of 29 candidates corresponded to previously-reported variants; 26 were novel predictions.

Conclusion. The study concluded that 55 prioritized nsSNPs were computationally predicted candidates for experimental validation, not confirmed pathogenic mutations. Biochemical follow-up may accelerate genotype–phenotype understanding in this.

Keywords: ferroportin (FPN), hemochromatosis type 4, *in silico* analysis, non-synonymous single nucleotide polymorphisms (nsSNPs), protein modelling, *SLC40A1*

*Corresponding Author: Misha.143@ucp.edu.pk

Highlights

- A total of 300 coding-region non-synonymous single nucleotide polymorphisms (nsSNPs) in SLC40A1 (ferroportin) were systematically screened *in silico* using four functional, two stability, and four structural prediction tools.
- Consensus-based filtering identified 55 high-confidence deleterious nsSNPs, comprising 29 previously reported variants and 26 novel predictions.
- Analysis using Project HOPE and I-TASSER indicated that 38 of the 55 prioritized variants may disrupt conserved transmembrane or functional domains of ferroportin.
- The prioritized variants were proposed as potential candidates for experimental validation through cellular iron-efflux and hepcidin-binding assays.

1. INTRODUCTION

Iron is an essential trace element that functions as an oxygen carrier in haemoglobin and myoglobin, as a cofactor of numerous enzymes, and as a component of the respiratory chain. Since mammals lack an active excretory pathway for iron, they tightly control whole-body iron at the level of intestinal absorption and macrophage recycling [1]. Dysregulation of this homeostasis underlies hereditary hemochromatosis (HH), a genetically heterogeneous disorder that features raised serum iron indices and progressive parenchymal iron deposition. This may cause hepatic cirrhosis, hepatocellular carcinoma, diabetes, arthropathy, and cardiomyopathy if patients do not receive treatment [2,3].

Hereditary hemochromatosis comprises four main types. Variants in HFE most commonly cause type 1 HH [4,5], whereas the non-HFE forms include juvenile HH types 2A (HJV) and 2B (HAMP), type 3 (TFR2) and type 4 (SLC40A1). Type 4 HH, also termed as ferroportin disease, is the only autosomal dominant form; variants in SLC40A1 cause it exclusively [6].

SLC40A1 at 2q32.2 encodes ferroportin (FPN; UniProt Q9NP59), a 571-amino-acid polytopic membrane protein of the major facilitator superfamily and the

only known mammalian cellular iron exporter. It is expressed most highly on the basolateral surface of duodenal enterocytes, on reticuloendothelial macrophages of liver and spleen, and on hepatocytes and placental syncytiotrophoblasts [6,7]. Recent cryo-electron microscopy structures place FPN in a canonical 12-transmembrane-helix fold organised into two six-helix bundles (N- and C-lobes) that alternate between outward- and inward-open conformations to translocate ferrous iron across the plasma membrane [8]. Plasma iron levels are controlled post-translationally by the liver-derived peptide hormone hepcidin. This binds the extracellular cavity of FPN, occludes the iron-permeation pathway as well as drives ubiquitin-dependent internalization and lysosomal degradation of the transporter [9,10].

Type 4 HH comprises two clinical entities with distinct underlying molecular mechanisms. Type 4A, or “classical” ferroportin disease, results from loss-of-function variants that impair FPN trafficking, membrane localization, or iron-efflux activity, leading to macrophage iron accumulation with normal-to-low transferrin saturation. In contrast, Type 4B, or “hepcidin-resistant” ferroportin disease, is caused by gain-of-function variants that preserve iron-export activity but disrupt the FPN–

hepcidin interaction, resulting in hepatocellular iron overload, elevated transferrin saturation, and hyperferritinaemia, with a phenotype resembling classical HFE-related HH [6,11,12]. Representative variants include the recurrent loss-of-function variant R178Q [13], the H507R variant reported in a Japanese case [14], the Italian D181N, G204R, and R296Q variants associated with variable defects in trafficking and hepcidin binding [15], and the gain-of-function C326F variant reported in a Chinese patient [16].

Despite the steady accumulation of catalogued SLC40A1 variants in public databases, the majority of reported nsSNPs have no functional annotation, and, to our knowledge, no study has reported a systematic triage of the full coding-region variant set. This creates an unmet need for a prioritised short-list of computationally-predicted high-confidence variants to guide experimental follow-up. The objective of the current study was therefore to apply an integrated *in silico* pipeline that combines sequence-based, structure-based, conservation-based, and stability-based predictors to

every coding-region nsSNP of *SLC40A1*, in order to identify and prioritise high-confidence deleterious candidates for experimental validation.

2. MATERIALS AND METHODS

2.1. Data Sources and Variant Retrieval

The study obtained the human SLC40A1 reference sequence and protein from NCBI (Gene ID 30061; RefSeq mRNA NM_014585.6; RefSeq protein NP_055400.1) and UniProt (accession Q9NP59, ferroportin; 571 aa). Furthermore, the study downloaded the complete catalogue of single-nucleotide variants annotated to the SLC40A1 coding region from NCBI dbSNP (build 156) and cross-verified it against Ensembl (release 110; transcript ENST00000261024). From this catalogue, only missense (non-synonymous) variants were retained present on the canonical transcript, yielding 300 nsSNPs for downstream analysis. All databases were accessed during the period in which the study (2022–2023) was performed.

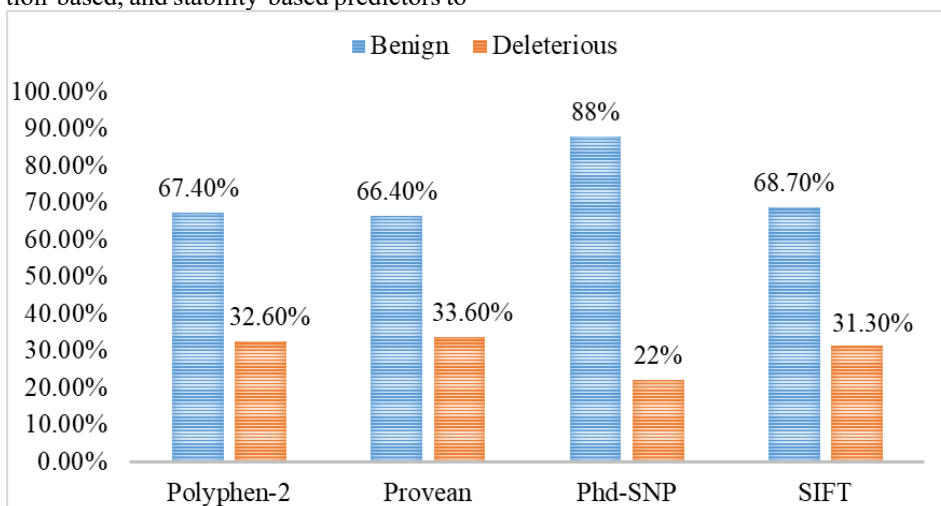


Figure 1. PRISMA-style Workflow Illustrating the *in silico* Screening and Prioritization Pipeline

2.2. Analysis Workflow

The full pipeline is summarised in Figure 1. Each of the 300 nsSNPs were evaluated sequentially by: (i) four functional-impact predictors, PolyPhen-2, PROVEAN, SIFT, and PhD-SNP; (ii) two thermodynamic-stability predictors, I-Mutant 3.0 and MUpro; (iii) residue-level structural analysis with Project HOPE; and (iv) secondary- and tertiary-structure analysis with PSIPRED, I-TASSER, and GalaxyRefine, with 3-D visualisation in Mutation3D and PyMOL. Variants classified as deleterious/disease-associated by all four functional predictors and destabilising ($\Delta\Delta G < 0$ or MUpro score < 0) by at least one of the two stability tools were retained as the high-confidence set. Thresholds, versions, and URLs for each tool are listed in Table 1 with a brief justification for the choice of predictor.

A total of 300 coding-region nsSNPs in *SLC40A1*, retrieved from NCBI dbSNP and Ensembl, were assessed using four functional predictors (PolyPhen-2, PROVEAN, SIFT, and PhD-SNP) and two stability predictors (I-Mutant 3.0 and MUpro). Variants predicted as deleterious by all four functional predictors and destabilising by at least one stability predictor were retained ($n = 55$) for subsequent residue-level, secondary-, tertiary-, and refined-structure analyses.

2.3. Tool Inventory, Versions, and Thresholds

Table 1 lists every predictor used in this study together with its version or access date, its URL, the decision threshold adopted for the present analysis, and a one-line rationale for its inclusion. The panel was chosen to provide complementary lines of evidence: PolyPhen-2 is structure-based, PROVEAN is alignment-based, SIFT is conservation-based, PhD-SNP is a support-

vector-machine (SVM) disease classifier, and I-Mutant 3.0 and MUpro provide orthogonal (structure- and sequence-based) thermodynamic-stability estimates. This combination reduces the chance of tool-specific false positives and false negatives.

2.4. Functional-impact Predictors

PolyPhen-2 (Polymorphism Phenotyping v2.2.3) ranks missense variants as benign, possibly damaging or probably damaging by combining sequence-based and structure-based features. The study adopted the HumDiv model and a score threshold of ≥ 0.85 for “probably damaging” [17]. PROVEAN (v1.1.3) scores a substitution as deleterious when the alignment-based score is ≤ -2.5 [18]. SIFT classifies a substitution as not tolerated when the score is ≤ 0.05 [19]. PhD-SNP uses an SVM classifier to assign each variant as disease-associated or neutral; the disease call requires a posterior probability > 0.5 [20].

2.5. Thermodynamic-stability Predictors

I-Mutant 3.0 predicts the change in folding free energy ($\Delta\Delta G$) induced by a single-site mutation on the basis of the protein structure. The study used the $\Delta\Delta G$ output in the sequence-only mode and treated any value < 0 kcal mol⁻¹ as destabilising [21]. MUpro is an orthogonal sequence-based SVM predictor of stability change; an output score < 0 was interpreted as destabilising [22].

2.6. Structural Analysis and Modelling

Residue-level structural consequences of each variant were assessed with Project HOPE. This integrates sequence annotation from UniProt, 3-D structural information, and physicochemical residue properties [23]. The secondary structure of ferroportin was predicted from the primary sequence with PSIPRED v4.0 using PSI-BLAST profiles [24]. Tertiary structure was modelled

with I-TASSER v5.1 [25]; of the five candidate models, the model with the highest C-score was retained. The selected model was refined with GalaxyRefine to improve

stereochemistry and MolProbity score [26], visualised in PyMOL, and used as the reference structure for the 55 high-confidence variants in Mutation3D [27].

Table 1. Computational Tools used, Versions, Decision Thresholds, and Rationale

Tool	Version / Access	URL	Deleterious Threshold	Rationale
PolyPhen-2	v2.2.3	http://genetics.bwh.harvard.edu/pph2/	HumDiv $\geq$ 0.85 (probably damaging)	Structure-based
PROVEAN	v1.1.3	http://provean.jcvi.org/	Score ≤ -2.5	Alignment-based
SIFT	SIFT Web v6.2.1	https://sift.bii.a-star.edu.sg/	Score ≤ 0.05 (not tolerated)	Sequence conservation
PhD-SNP	Web server, accessed 2023	https://snps.biofold.org/phd-snp/	Probability > 0.5 (disease)	SVM disease classifier
I-Mutant 3.0	v3.0	http://gpcr2.biocomp.unibo.it/cgi/predictors/I-Mutant3.0/	$\Delta\Delta G < 0$ (destabilizing)	SVM stability (structure)
MUpro	Web server, accessed 2023	http://mupro.proteomics.ics.uci.edu/	Score < 0 (destabilizing)	SVM stability (sequence)
HOPE	Web server, accessed 2023	https://www3.cmbi.umcn.nl/hope/	Qualitative impact report	Residue-level structure effect
PSIPRED	v4.0	http://bioinf.cs.ucl.ac.uk/psipred/	n/a	Secondary structure
I-TASSER	v5.1	https://zhanggroup.org/I-TASSER/	Highest C-score model	Tertiary structure
GalaxyRefine	Web server, accessed 2023	http://galaxy.seoklab.org/	Improved MolProbity score	Structure refinement
Mutation3D	Web server, accessed 2023	http://mutation3d.org/	Qualitative	3D visualization of variants

2.7. Consensus Criterion for Prioritization

The study classified a variant as a high-confidence deleterious candidate only if all four functional predictors (PolyPhen-2, PROVEAN, SIFT and PhD-SNP) called it deleterious or disease-associated and at least one of the two stability predictors called it destabilising (score < 0 or $\Delta\Delta G < 0$). Application of this criterion to the 300 input nsSNPs yielded the final set of 55 variants reported in Section 3.

3. RESULTS

3.1. Input Variant Set

The current study retrieved a total of 300 coding-region nsSNPs in SLC40A1 from NCBI dbSNP (build 156) and cross-verified them in Ensembl (release 110) on the canonical transcript ENST00000261024. This set formed the input to the downstream analysis pipeline (Figure 1).

3.2. Functional-impact Predictions

PolyPhen-2 classified 98 of 300 variants (32.6%) as probably damaging and the

remaining 202 (67.4%) as benign or possibly damaging. PROVEAN classified 101 (33.6%) as deleterious and 199 (66.4%) as neutral. SIFT classified 94 (31.3%) as not tolerated and 206 (68.7%) as tolerated. PhD-SNP classified 66 variants (22.0%) as disease-associated and 234 (78.0%) as neutral. These proportions correct an inconsistency present in the previous version of the manuscript, in which the disease and neutral fractions did not sum to 100%. Agreement across all four functional predictors ("call deleterious" by each) was achieved for 62 variants (Figure 2).

Bars show the percentage of the 300 input nsSNPs classified as deleterious (damaging / disease-associated / not tolerated) by PolyPhen-2, PROVEAN, SIFT and PhD-SNP, respectively. The updated PhD-SNP proportions (22.0% disease-associated and 78.0% neutral) correct the numerical inconsistency reported in the previous version.

3.3. Thermodynamic-stability Predictions

Of the 300 input variants, I-Mutant 3.0

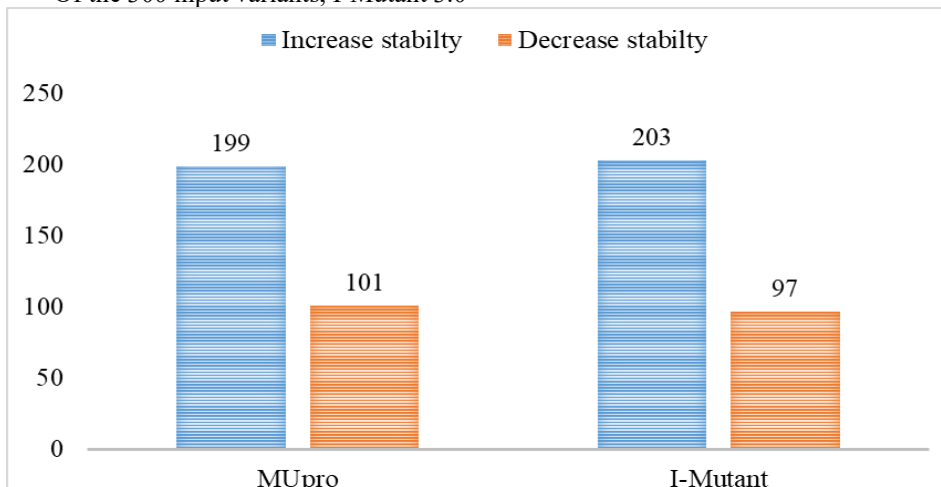


Figure 2. Summary of Pathogenicity Predictions across the Four Functional Predictors.

predicted 97 (32.3%) to reduce folding stability ($\Delta\Delta G < 0$) and 203 (67.7%) to increase or preserve stability. MUpro produced concordant results, classifying 101 (33.7%) as destabilising and 199 (66.3%) as stabilising or neutral. Agreement between the two stability predictors was observed for the large majority of the variants that were also deleterious by all four functional predictors.

3.4. Prioritized High-Confidence nsSNPs

Application of the consensus criterion (Section 2.7) yielded 55 high-confidence deleterious nsSNPs. Table 2 lists each variant with its predicted domain localisation on ferroportin, the call from each predictor, and a novel/previously-reported flag based on cross-referencing with the published *SLC40A1* literature and ClinVar. Of the 55 candidates, 29 corresponded to previously-reported *SLC40A1* variants (several with experimental support), and 26 were novel computational predictions that, to our knowledge, have not previously been characterized (Figure 3).

Table 2. Summary of the 55 High-confidence nsSNPs of SLC40A1 Prioritised by the Consensus Pipeline

#	AA change	Predicted domain/location	Poly-Phen-2	PROV EAN	SIFT	PhD-SNP	I-Mutant	MUpro	Status
1	L30P	N-terminal cytoplasmic	PD	D	NT	Dis	Dec	Dec	Novel
2	R40W	TM1	PD	D	NT	Dis	Dec	Dec	Novel
3	M41I	TM1	PD	D	NT	Dis	Dec	Dec	Novel
4	Y54C	TM1	PD	D	NT	Dis	Dec	Dec	Reported
5	S57R	TM1	PD	D	NT	Dis	Dec	Dec	Novel
6	L60W	TM1	PD	D	NT	Dis	Dec	Dec	Novel
7	T61I	TM1	PD	D	NT	Dis	Dec	Inc	Reported
8	Y64N	TM1-2 loop	PD	D	NT	Dis	Dec	Dec	Novel
9	Y64C	TM1-2 loop	PD	D	NT	Dis	Dec	Dec	Novel
10	G65R	TM1-2 loop	PD	D	NT	Dis	Dec	Dec	Novel
11	L75P	TM2	PD	D	NT	Dis	Dec	Dec	Reported
12	A77D	TM2	PD	D	NT	Dis	Dec	Dec	Novel
13	G80S	TM2	PD	D	NT	Dis	Dec	Dec	Novel
14	G80V	TM2	PD	D	NT	Dis	Dec	Dec	Reported
15	W82G	TM2	PD	D	NT	Dis	Dec	Dec	Novel
16	V83G	TM2	PD	D	NT	Dis	Dec	Dec	Novel
17	R88G	TM2-3 loop	PD	D	NT	Dis	Dec	Dec	Reported
18	R88T	TM2-3 loop	PD	D	NT	Dis	Dec	Dec	Reported
19	L89R	TM2-3 loop	PD	D	NT	Dis	Dec	Dec	Reported
20	L110R	TM3	PD	D	NT	Dis	Dec	Dec	Reported
21	N144H	TM4	PD	D	NT	Dis	Dec	Dec	Reported
22	N144Y	TM4	PD	D	NT	Dis	Dec	Dec	Novel
23	I152F	TM4	PD	D	NT	Dis	Dec	Dec	Reported
24	D157N	TM4	PD	D	NT	Dis	Dec	Dec	Reported
25	D157G	TM4	PD	D	NT	Dis	Dec	Dec	Reported
26	D157V	TM4	PD	D	NT	Dis	Dec	Dec	Novel
27	W158L	TM4	PD	D	NT	Dis	Dec	Dec	Reported
28	W158C	TM4	PD	D	NT	Dis	Dec	Dec	Reported
29	A175D	TM4-5 loop	PD	D	NT	Dis	Dec	Dec	Reported
30	R178G	TM4-5 loop	PD	D	NT	Dis	Dec	Dec	Reported
31	R178Q	TM4-5 loop	PD	D	NT	Dis	Dec	Dec	Reported
32	R179T	TM4-5 loop	PD	D	NT	Dis	Dec	Dec	Novel
33	I180T	TM4-5 loop	PD	D	NT	Dis	Dec	Dec	Reported
34	D181N	TM5	PD	D	NT	Dis	Dec	Dec	Reported
35	D181V	TM5	PD	D	NT	Dis	Dec	Dec	Reported
36	S200Y	TM5	PD	D	NT	Dis	Dec	Dec	Reported
37	S200F	TM5	PD	D	NT	Dis	Dec	Dec	Reported
38	G204V	Intracellular loop	PD	D	NT	Dis	Dec	Dec	Reported
39	C205W	Intracellular loop	PD	D	NT	Dis	Dec	Dec	Novel
40	L233P	TM6	PD	D	NT	Dis	Dec	Dec	Reported
41	W299C	TM7	PD	D	NT	Dis	Dec	Dec	Reported
42	T320I	TM7-8 loop	PD	D	NT	Dis	Dec	Dec	Novel
43	F324S	TM8	PD	D	NT	Dis	Dec	Dec	Reported
44	S338R	TM8	PD	D	NT	Dis	Dec	Dec	Reported
45	A350D	TM8-9 loop	PD	D	NT	Dis	Dec	Dec	Reported
46	W362S	TM9	PD	D	NT	Dis	Dec	Dec	Novel
47	W362C	TM9	PD	D	NT	Dis	Dec	Dec	Novel

#	AA change	Predicted do-main/location	Poly-Phen-2	PROV EAN	SIFT	PhD-SNP	I-Mu-tant	MUpro	Status
48	R364C	TM9	PD	D	NT	Dis	Dec	Dec	Novel
49	C367R	TM9	PD	D	NT	Dis	Dec	Dec	Novel
50	V370G	TM9	PD	D	NT	Dis	Dec	Dec	Reported
51	R371G	TM9	PD	D	NT	Dis	Dec	Dec	Novel
52	R371W	TM9	PD	D	NT	Dis	Dec	Dec	Novel
53	G373S	TM9	PD	D	NT	Dis	Dec	Dec	Reported
54	L384F	TM9-10 loop	PD	D	NT	Dis	Dec	Dec	Novel
55	S390F	TM10	PD	D	NT	Dis	Dec	Dec	Reported

Note. PD = probably damaging; D = deleterious; NT = not tolerated; Dis = disease-associated; Dec = decrease in stability; Inc = increase in stability.



Figure 3. Protein-stability Predictions for the 300 nsSNPs from MUpro and I-Mutant 3.0

3.5. Residue-level Structural Consequences

Project HOPE analyses of the 55 high-confidence variants indicated that 38 substitutions were predicted to produce substantial structural perturbation of ferroportin. Whereas, 17 were predicted to produce more subtle effects. The 38 high-impact substitutions clustered in the transmembrane helices and in the TM4–5 loop that lines the extracellular hepcidin-binding cavity. Mechanistically relevant examples include: Y54C and W158C, in which introduction of a free cysteine into a hydrophobic transmembrane environment is pre-

dicted to disrupt helical packing and potentially introduce aberrant disulphide bonds; D157N and D181N, in which loss of a buried acidic side chain is expected to perturb the iron-coordination network; R178Q, which removes a positive charge implicated in the iron-permeation pathway and is consistent with the loss-of-function phenotype previously reported experimentally [13]; and C326F in TM8 (reported elsewhere [16]), which has been associated with gain-of-function/hepcidin-resistance. Taken together, these residue-level predictions are mechanistically coherent with the two clinical subtypes of ferroportin disease (4A versus 4B).



Figure 4. Secondary Structure of Ferroportin Predicted by PSIPRED from the 571-residue Primary Sequence. Helical (H), strand (E), and coil (C) elements are Shown Together with the PSIPRED Confidence Track

Table 3. Comparison of the Tertiary Models of Ferroportin Generated by I-TASSER and Refined by GalaxyRefine

Model	GDT-HA	RMSD	MolPro-bity	Clash score	Poor rotamers	Rama favoured
Model produced by I-TASSER	1.0000	0.000	3.099	12.2	10.7	81.4
GalaxyRefine Model 1	0.9440	0.482	2.283	19.6	0.2	91.7

3.6. Secondary and Tertiary Structure of Ferroportin

PSIPRED predicted the 571-residue ferroportin sequence to fold into approximately 12 α -helical transmembrane segments with connecting intra- and extracellular loops and short N- and C-terminal cytoplasmic extensions (Figure 4), consistent with the major-facilitator-superfamily fold reported in recent cryo-EM structures [8]. The I-TASSER tertiary model with the highest C-score was selected, refined with GalaxyRefine and visualised in PyMOL (Figure 5). Mutant models for each of the 55 high-confidence variants are provided as Supplementary Figures S1–S55.

**Figure 5.** Tertiary Structure of Ferroportin.

The I-TASSER model with the highest C-score was refined with GalaxyRefine and rendered in PyMOL. The 12 transmembrane helices organise into N- and C-terminal six-helix bundles consistent with the major-facilitator-superfamily fold.

3.7. Structure Refinement

Comparison of the raw I-TASSER model and the GalaxyRefine-refined Model

1 (Table 3) showed that refinement improved the MolProbity score from 3.099 to 2.283. Moreover, it raised the proportion of residues in the Ramachandran-favored region from 81.4% to 91.7%, with only a minor increase in clash score. The study therefore adopted the refined model for all downstream structural analyses.

4. DISCUSSION

Hemochromatosis type 4 is the only autosomal dominant form of hereditary hemochromatosis; unlike the HFE-related forms, variants in a single gene, *SLC40A1*, cause it exclusively [5,6]. Ferroportin, the protein product of *SLC40A1*, is the only characterised cellular iron exporter in mammals and is post-translationally controlled by hepcidin-mediated internalisation [9,10]. Since the two clinical sub-types of ferroportin disease (4A, loss of function; 4B, hepcidin resistance) arise from mechanistically distinct classes of missense variants, residue-level annotation of the full coding-region nsSNP catalogue is a prerequisite for accurate genotype–phenotype correlation.

The current study evaluated 300 coding-region nsSNPs using a combination of functional-impact, thermodynamic-stability, and structural predictors. The application of a conservative consensus criterion yielded 55 high-confidence deleterious candidates. Of these, 29 corresponded to previously catalogued or published *SLC40A1* variants. Reassuringly, our pipe-

line recovered variants with direct experimental support: R178Q has been shown biochemically to cause loss of iron-export function [13]; W158C has been associated with reduced iron export [28]; D157G has been linked to hepcidin-independent Jak2-mediated down-regulation of ferroportin [29]; G204R produces a complex functional defect affecting both iron transport and hepcidin sensitivity [15]; and C326F underlies a gain-of-function/hepcidin-resistant phenotype [16]. Recovery of these experimentally validated variants supports the construct validity of our pipeline.

Beyond these “positive controls”, 26 of the 55 candidates are novel computational predictions (Table 2), including N144Y, R179T, C205W, T320I, W362S, W362C, R364C, C367R, R371G, R371W, and L384F. Several of these variants lie within or immediately adjacent to transmembrane helices. Disruption of those helices would compromise the alternating-access mechanism of ferroportin [8], and two (R371G and R371W) affect a basic residue that may participate in the hepcidin-binding cavity. These variants, therefore, represent mechanistically plausible candidates for future experimental validation.

The results of this study should be interpreted in the context of the inherent limitations of *in silico* variant analysis. Firstly, all predictions presented here are computational hypotheses; they do not replace biochemical, cellular or clinical evidence. Computational predictors are trained on non-random subsets of known variants and can therefore propagate training-set biases, particularly for membrane proteins. Secondly, thermodynamic-stability predictors, such as I-Mutant 3.0 and MUpro, estimate globular-fold stability and may underestimate the effect of substitutions in lipid-embedded transmembrane segments. Thirdly, the current study did not incorporate allele

frequencies (for example, from gnomAD) or segregation data from affected kindreds, which would further refine clinical interpretation. Finally, functional consequences of variants in ferroportin can be category-specific (loss of iron transport, loss of surface expression, or loss of hepcidin responsiveness). Additionally, sequence-based predictors alone cannot resolve these distinct phenotypes. For all of these reasons, the 55 candidates presented here should be regarded as a prioritised list for experimental follow-up in iron-efflux assays, surface-biotinylation assays, and hepcidin-binding assays, rather than as confirmed pathogenic mutations.

With these caveats explicit, the present integrated pipeline achieves what single-tool analyses cannot: it filters a large nsSNP catalogue through orthogonal evidence streams, reduces tool-specific false positives, and produces a manageable list of mechanistically plausible variants concentrated in the conserved transmembrane core of ferroportin. The 26 novel predictions are the most immediate experimental priority, and the 29 previously-reported variants provide internal concordance with the published literature.

4.1. Conclusion

Using an integrated *in silico* pipeline that combines functional-impact, stability, and structural predictors, the study prioritised 55 high-confidence candidate deleterious nsSNPs of *SLC40A1* from a starting set of 300 coding-region variants. Twenty-nine of the 55 candidates corresponded to previously-reported *SLC40A1* variants, providing internal validation of the pipeline; the remaining 26 were novel computational predictions. These variants should be regarded as candidates for further experimental validation, for instance, in iron-efflux, surface-expression, and hepcidin-

binding assays, and not as confirmed pathogenic mutations. With appropriate biochemical and clinical follow-up, the prioritised set reported here may accelerate the identification of genotype–phenotype relationships and of potential therapeutic targets in ferroportin disease.

Author Contribution

Rimsha Zafar: Writing – Original Draft, Data Curation. **Naveera Khan:** Formal Analysis. **Muhammad Awaiz:** Supervision. **Zartash Butt:** Writing – Review & Editing.

Conflict of Interest

The authors of the manuscript have no financial or non-financial conflict of interest in the subject matter or materials discussed in this manuscript.

Data Availability Statement

The source SLC40A1 nsSNP data are publicly available through NCBI dbSNP build 156 (Gene ID 30061), Ensembl release 110 (transcript ENST00000261024), and UniProt (accession Q9NP59). The mutant structural models generated using I-TASSER and Mutation3D are available from the corresponding author upon reasonable request.

Funding Details

No funding has been received for this research.

Generative AI Disclosure Statement

The authors did not use any type of generative artificial intelligence software for this research.

REFERENCES

1. Kowdley DS, Kowdley KV. Appropriate clinical genetic testing of hemochromatosis type 2–4, including ferroportin disease. *Appl Clin Genet*. 2021;14:353–361. <https://doi.org/10.2147/TACG.S269622>
2. Girelli D, Busti F, Brissot P, Cabantchik I, Muckenthaler MU, Porto G. Hemochromatosis classification: update and recommendations by the BIIROIRON Society. *Blood*. 2022;139(20):3018–3029. <https://doi.org/10.1182/blood.2021011338>
3. Brissot P, Cavey T, Ropert M, Gugenbuhl P, Loréal O. Genetic hemochromatosis: pathophysiology, diagnostic and therapeutic management. *Presse Med*. 2017;46(12 Pt 2):e288–e295. <https://doi.org/10.1016/j.lpm.2017.05.037>
4. Ferbo L, Manzini PM, Badar S, et al. Detection of a rare mutation in the ferroportin gene through targeted next generation sequencing. *Blood Transfus*. 2016;14(6):531–534. <https://doi.org/10.2450/2016.0286-15>
5. Pietrangelo A. Hereditary hemochromatosis: pathogenesis, diagnosis, and treatment. *Gastroenterology*. 2010;139(2):393–408.e2. <https://doi.org/10.1053/j.gastro.2010.06.013>
6. Choi EK, Nguyen TT, Iwase S, Seo YA. Ferroportin disease mutations influence manganese accumulation and cytotoxicity. *FASEB J*. 2019;33(2):2228–2240. <https://doi.org/10.1096/fj.201800831R>
7. Qiao B, Sugianto P, Fung E, et al. Hepcidin-induced endocytosis of ferroportin is dependent on ferroportin ubiquitination. *Cell Metab*. 2012;15(6):918–924. <https://doi.org/10.1016/j.cmet.2012.03.018>
8. Ka C, Guellec J, Pepermans X, et al. The SLC40A1 R178Q mutation is a recurrent cause of hemochromatosis and is associated with a novel pathogenic mechanism. *Haematologica*. 2018;103(11):1796–1805. <https://doi.org/10.3324/haematol.2018.189845>
9. Yamakawa N, Oe K, Yukawa N, et al.

- A novel phenotype of hereditary hemochromatosis type 4 with ferroportin-1 mutation, presenting with juvenile cataracts. *Intern Med*. 2016;55(18):2697–2701. <https://doi.org/10.2169/internalmedicine.55.6565>
10. Majore S, Di Patti MCB, Valiante M, et al. Characterization of three novel pathogenic SLC40A1 mutations and genotype/phenotype correlations in 7 Italian families with type 4 hereditary hemochromatosis. *Biochim Biophys Acta Mol Basis Dis*. 2018;1864(2):464–470. <https://doi.org/10.1016/j.bbadis.2017.11.006>
 11. Chen SR, Yang LQ, Chong YT, et al. Novel gain-of-function mutation in the SLC40A1 gene associated with hereditary haemochromatosis type 4. *Intern Med J*. 2015;45(6):672–676. <https://doi.org/10.1111/imj.12764>
 12. Awais M, Wattoo JI, Zafar R, Khan N. Computational analysis of non-synonymous single nucleotide polymorphism in UROD gene linked with familial porphyria cutanea tarda. *Gene Rep*. 2021;25:101413. <https://doi.org/10.1016/j.genrep.2021.101413>
 13. Adzhubei I, Jordan DM, Sunyaev SR. Predicting functional effect of human missense mutations using PolyPhen-2. *Curr Protoc Hum Genet*. 2013;76(1):7.20.1–7.20.41. <https://doi.org/10.1002/0471142905.hg0720s76>
 14. Choi Y, Chan AP. PROVEAN web server: a tool to predict the functional effect of amino acid substitutions and indels. *Bioinformatics*. 2015;31(16):2745–2747. <https://doi.org/10.1093/bioinformatics/btv195>
 15. Sim NL, Kumar P, Hu J, Henikoff S, Schneider G, Ng PC. SIFT web server: predicting effects of amino acid substitutions on proteins. *Nucleic Acids Res*. 2012;40(W1):W452–W457. <https://doi.org/10.1093/nar/gks539>
 16. Kumar P, Henikoff S, Ng PC. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nat Protoc*. 2009;4(7):1073–1081. <https://doi.org/10.1038/nprot.2009.86>
 17. Capriotti E, Fariselli P. PhD-SNPg: a webserver and lightweight tool for scoring single nucleotide variants. *Nucleic Acids Res*. 2017;45(W1):W247–W252. <https://doi.org/10.1093/nar/gkx369>
 18. Cheng J, Randall A, Baldi P. Prediction of protein stability changes for single-site mutations using support vector machines. *Proteins*. 2006;62(4):1125–1132. <https://doi.org/10.1002/prot.20810>
 19. Kepp KP. Towards a “golden standard” for computing globin stability: stability and structure sensitivity of myoglobin mutants. *Biochim Biophys Acta Proteins Proteom*. 2015;1854(10):1239–1248. <https://doi.org/10.1016/j.bbapap.2015.06.002>
 20. Venselaar H, te Beek TAH, Kuipers RKP, Hekkelman ML, Vriend G. Protein structure analysis of mutations causing inheritable diseases: an e-Science approach with life scientist friendly interfaces. *BMC Bioinformatics*. 2010;11:548. <https://doi.org/10.1186/1471-2105-11-548>
 21. Buchan DWA, Minneci F, Nugent TCO, Bryson K, Jones DT. Scalable web services for the PSIPRED Protein Analysis Workbench. *Nucleic Acids Res*. 2013;41(W1):W349–W357. <https://doi.org/10.1093/nar/gkt381>

22. Yang J, Yan R, Roy A, Xu D, Poisson J, Zhang Y. The I-TASSER Suite: protein structure and function prediction. *Nat Methods*. 2015;12(1):7–8. <https://doi.org/10.1038/nmeth.3213>
23. Meyer MJ, Lapcevic R, Romero AE, et al. mutation3D: cancer gene prediction through atomic clustering of coding variants in the structural proteome. *Hum Mutat*. 2016;37(5):447–456. <https://doi.org/10.1002/humu.22963>
24. Ko J, Park H, Heo L, Seok C. Galaxy-WEB server for protein structure prediction and refinement. *Nucleic Acids Res*. 2012;40(W1):W294–W297. <https://doi.org/10.1093/nar/gks493>
25. Ekanayake D, Roddick C, Powell LW. Recent advances in hemochromatosis: a 2015 update: a summary of proceedings of the 2014 conference held under the auspices of Hemochromatosis Australia. *Hepatol Int*. 2015;9(2):174–182. <https://doi.org/10.1007/s12072-015-9608-2>
26. Abboud S, Haile DJ. A novel mammalian iron-regulated protein involved in intracellular iron metabolism. *J Biol Chem*. 2000;275(26):19906–19912. <https://doi.org/10.1074/jbc.M000713200>
27. Mayr R, Janecke AR, Schranz M, et al. Ferroportin disease: a systematic meta-analysis of clinical and molecular findings. *J Hepatol*. 2010;53(5):941–949. <https://doi.org/10.1016/j.jhep.2010.05.016>
28. Wallace DF, Subramaniam VN. Non-HFE hemochromatosis. In: Anderson GJ, McLaren GD, eds. *Iron Physiology and Pathophysiology in Humans*. Springer; 2012:399–416. https://doi.org/10.1007/978-1-60327-485-2_20
29. Pietrangelo A. Ferroportin disease: pathogenesis, diagnosis and treatment. *Haematologica*. 2017;102(12):1972–1984. <https://doi.org/10.3324/haematol.2017.170720>