In-silico predictions of deleterious SNPs in Human ephrin type-A receptor 3 (EPHA3) gene

Dipankor Chatterjee

University of Dhaka Faculty of Biological Sciences

Umar Faruq Chowdhury (✉ umarfchy@gmail.com)

University of Dhaka Faculty of Biological Sciences <https://orcid.org/0000-0001-5088-4658>

Mohammad Umer Sharif Shohan

University of Dhaka Faculty of Biological Sciences

Md Mohasin

University of Dhaka Faculty of Biological Sciences

Yearul Kabir

University of Dhaka Faculty of Biological Sciences

Research Article

Keywords: Single nucleotide polymorphism (SNP), non-synonymous SNPs (nsSNP), Ephrin type-A receptor 3, Adenocarcinoma, dbSNP database

Posted Date: April 22nd, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-356150/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Version of Record: A version of this preprint was published at Informatics in Medicine Unlocked on January 1st, 2021. See the published version at <https://doi.org/10.1016/j.imu.2021.100728>.

Abstract

Ephrin type-A receptor 3 (EPHA3) is a receptor tyrosine kinase encoded by the EPHA3 gene, involves in many biological functions, such as cell migration, adhesion, etc. Overexpression or inactivation of EPHA3 receptors can be complicit in various oncogenic events. Therefore, SNP analysis is required to screen out the deleterious SNPs, such as non-synonymous SNPs (nsSNPs), and *in-silico* analysis is an effective way to do this screening process. Using eight bioinformatics tools, 81 nsSNPs were found as potential harmful nsSNPs by *in-silico* analysis among 631 nsSNPs that were retrieved from the dbSNP database. 77 out of 81 substitutions were found as highly conserved by ConSurf analysis. G628E, I652M nsSNPs were found to be present in the ATP-binding site. N751D was found in one of the catalytic residues and P824S in the α F- α G loop in the substrate-binding region. Y596H, Y779C & D598G nsSNPs were found to be associated with activation and regulatory mechanism. R66S, R160H, R104Q nsSNPs were present in the EPH-ligand binding domain that has a higher probability of interfering with ligand-receptor interactions. 21 SNPs associated with the alteration of miRNA target site at 3' UTR region, which may be responsible for affecting post-transcriptional regulation of EPHA3 gene. These findings will provide necessary data to explore various diseases or cancers pertinent to the EPHA3 receptor protein, will provide remedial biomarkers, may help in molecular diagnosis, and help to design target-specific therapeutic agents.

Introduction

The ephrin receptor (EPH) family, one of the largest sub-group of receptor tyrosine kinase, involves many important functions, such as development, directing segmentation, axon guidance, fasciculation, angiogenesis, and limb development [1]. These proteins contain five domains, such as an extracellular ligand domain, two types III fibronectin domains, and two intracellular domains: the tyrosine kinase and sterile-alpha-motif (SAM) domains. Two types of ephrins (EPH ligands), type A ephrins, type B ephrins, tightly interact with the proteins through a glycosyl-phosphatidylinositol (GPI) moiety or a transmembrane domain. The interaction between the Eph receptor and ephrin-containing cells lead to bidirectional signaling, the events led to a series that either stabilize cell-cell contact or lead to cell repulsion. Eph receptors involve in many pathways, like cell shape via Rho activation, cell growth via the MAPK pathway, cell-cell adhesion via cadherin complexes, and cell-matrix interactions via integrin complexes by the downstream signaling from EPH receptors. Receptor-ligand dimers form heterotetramers, which further assemble into higher-order signaling clusters. A high-affinity binding site in the N-terminal domain of EPH receptor mediates intercellular Eph-ephrin interaction. Intracellular signaling by EphA3 receptors is initiated by autophosphorylation of three defined tyrosine residues, two in the highly conserved juxtamembrane region and the third in the activation loop of the kinase domain (Y779) [2]. Rapid reorganization of the actin and myosin cytoskeleton follows, leading to retraction of cellular protrusions, membrane blebbing, and cell detachment, followed by the association of the adaptor protein CrkII with tyrosine-phosphorylated EphA3 and activation of RhoA signaling [3]. So, EPHA3 is a vital protein receptor for biological activities.

EPHA3 has been seen to be highly expressed during vertebrate development. EPHA3 expression is restricted in adult tissues and detected significantly at a low level compared to the early developmental stage of vertebrates. Abnormal expression of EPHA3 is associated with many cancers. These abnormalities may occur due to mutational changes in specific regions of the EPHA3 gene. So EPHA3 receptor protein analysis should be conducted to better understand the diseases pertaining to the receptor as well as to strategize a therapeutic regime. By now, many studies reported that this protein is associated with many cancers, such as lung adenocarcinoma, colorectal carcinoma, metastatic melanoma, pancreatic cancer, and many others.

In-silico studies are now a very popular approach to find specific point mutations of a particular protein that has been associated with diseases or cancers. Many tools have been used in these *in-silico* based approaches to screening out the functional deleterious SNPs of a particular gene. And these tools can predict any changes associated with structure and functions and then provide scores based on their algorithms, which have quite a high degree of accuracy. By using these *in-silico* methods, functional and structural SNPs of many genes, such as BRCA1 (Breast cancer type 1 susceptibility protein), BRAF (B-Raf proto-oncogene) [4], IGF1R (Insulin-like growth factor 1 receptor) [5], have been identified.

In this study, several computational approaches have been carried out to find the most deleterious nsSNPs in the EPHA3 gene, which may be associated with dysfunctionality, structural instability, and are susceptible to many diseases or cancers. This study will help to extend the knowledge related to the effect of nsSNPs on protein functions and the stability of the EPHA3 protein.

Materials And Methods

Collection of nsSNPs

SNPs for the EPHA3 gene were collected from the NCBI dbSNP database [6] (<http://www.ncbi.nlm.nih.gov/SNP/>) and EPHA3 protein sequence (FASTA format) was collected from the Uniprot database (<https://www.uniprot.org/>). SNPs related to *Homo sapiens* were taken for analysis. Selected SNPs associated with non-synonymous, missense, non-sense were being filtered out for analysis.

Identification of damaging nsSNPs by SIFT

Non-synonymous SNPs (nsSNPs) of the EPHA3 gene were first screened through Sort Intolerant From Tolerant (SIFT, <https://sift.bii.a-star.edu.sg/>) program for finding out the most damaging SNPs. SIFT can predict if an amino acid substitution affects protein functions or not by analyzing sequence homology with other similar proteins and the physical properties of amino acids. Prediction scores less than or equal to 0.05 are considered as damaging and greater than 0.05 are considered to be tolerated [7].

Analyzing the effects of damaging SNPs

To further solidify the SIFT predictions, PolyPhen-2, I-Mutant Suite, and PROVEAN tools were used. PolyPhen-2 (Phenotyping Polymorphism version 2) tool can predict the effects of nsSNPs on structure and functions of the proteins by using machine learning method, TMHMM (transmembrane helix prediction by hidden markov model) algorithm, and high-quality multiple sequence alignment. PolyPhen-2 predicts the effects of nsSNPs and categorizes nsSNPs into probably damaging, possibly damaging, and benign SNPs [8].

I-Mutant Suite predicts changes in protein stability for any single point mutation by using a Support Vector Machine (SVM) based predictor [9].

PROVEAN (Protein Variation Effect Analyzer, <http://provean.jcvi.org/index.php>) tool which can predict any impact on protein functions due to substitution or insertion-deletion of any amino acid, were used. [10].

SNPs predicted from the above tools were further subjected to four different bioinformatics tools, SNPs&GO [11], SNAP-2 [12], PON-P2 [13], Mutant assessor [14]. As these tools use different algorithms, the predicted SNPs were further selected for downstream analysis.

Identification of nsSNPs in protein domains

A FASTA sequence of EPHA3 protein was submitted to the InterPro server (<https://www.ebi.ac.uk/interpro/>). InterPro server predicts families, conserved domains, and important sites by using signatures provided by many databases. Positions of nsSNPs in different domains can be predicted through this server [15].

EPHA3-protein interactions by STRING database:

STRING v11.0 (<https://string-db.org/>) was used for studying protein-protein interactions of EPHA3 with a high confidence score ($\geq 90\%$). The server identifies functional and physical associations by computational predictions. It also connects other databases to better understand network interactions of proteins [16].

Conservational analysis of EPHA3

ConSurf bioinformatics tool (<https://consurf.tau.ac.il/>) is used for the analysis of evolutionary conservation of nsSNPs positions in a protein sequence [17]. In this tool, phylogenetic relations are being analyzed between homologous proteins. The FASTA sequence of EPHA3 protein was submitted to the server, and it screened out highly conserved nsSNPs [18].

Besides, phylogenetic analysis of human EPHA3 protein sequence (UniProtKB ID EPHA3_HUMAN) and eleven different species protein sequences were collected from UniProt and analyzed for evolutionary conservation. Those species were *Pan troglodytes* (UniProtKB ID H2QMY7_PANTR), *Ictidomys tridecemlineatus* (UniProtKB ID I3MNR9_ICTTR), *Ictalurus punctatus* (UniProtKB ID W5UFX2_ICTPU), *Oreochromis niloticus* (UniProtKB ID I3J961_ORENI), *Macaca fascicularis* (UniProtKB ID G7NZ61_MACFA), *Mus musculus* (UniProtKB ID Q8BRB1_MOUSE), *Bos taurus* (UniProtKB ID

E1BJS9_BOVIN), *Gallus gallus* (UniProtKB ID F1NJX6_CHICK), *Rattus norvegicus* (UniProtKB ID G3V9D5_RAT), *Equus caballus* (UniProtKB ID F7C3U9_HORSE) and *Ovis aries* (UniProtKB ID W5Q6R5_SHEEP). Multiple sequence alignment was done by ClustalW and a phylogenetic tree was generated, taking 1000 bootstraps in MEGA-X [19].

Identification of Functional nsSNPs in 3'UTR by PolymiRTS

PolymiRTS (Polymorphism in microRNAs and their target sites, <http://compbio.uthsc.edu/miRSNP/>) database comprises naturally occurring DNA variations in microRNA (miRNA) seed regions and miRNA target sites [20]. PolymiRTS gives results in four forms D, N, C, and O, which indicates the derived allele disrupts a conserved miRNA site, the derived allele disrupts a nonconserved miRNA site, the derived allele creates a new miRNA site, and other cases respectively.

Results

The progression of the study, methods, and tools that were used to find damaging nsSNPs for the EPHA3 gene were summarized in Fig. 1.

Retrieval of SNPs from dbSNP database

EPHA3 gene information was collected from the NCBI dbSNP database. The database contains 81907 SNPs regarding EPHA3 gene, among which 631 (0.77%) SNPs are nsSNPs, 79839 (97.48%) SNPs are intron SNPs, 287 (0.35%) SNPs are coding synonymous SNPs, 42 (0.05%) SNPs are 5' UTR SNPs, 793 (0.97%) SNPs are 3' UTR SNPs and 315 (0.38%) SNPs are other types of SNPs (Fig. 2). nsSNPs on the EPHA3 gene were taken for analysis as they can alter amino acids.

Deleterious nsSNPs identified by SIFT

SIFT predicts damaging SNPs by analyzing sequence homology of the homologous proteins and physical properties of amino acids. Out of 631 nsSNPs, SIFT predicted 281 nsSNPs as damaging nsSNPs, and these SNPs were further analyzed in the rest of the tools.

Analysis of the effects of deleterious nsSNPs

PolyPhen-2 predicts probably damaging nsSNPs as the most damaging nsSNPs with high confidence prediction, and out of 281 nsSNPs, PolyPhen-2 predicted 222 nsSNPs to be most damaging. I-Mutant 3.0 (Suite) predicted 183 nsSNPs associated with the alteration of the stability of the protein. PROVEAN algorithm predicted 211 nsSNPs as deleterious nsSNPs that could affect protein functions.

SNPs&GO predicted 165 nsSNPs as deleterious nsSNPs, SNAP-2 predicted 211 affected nsSNPs, PON-P2 predicted 221 pathological nsSNPs and Mutant assessor predicted 183 nsSNPs as disease-associated nsSNPs (Supplementary Table 2). The predictions are listed in Supplementary Table 1.

After analyzing all eight different computational approaches, 81 nsSNPs were found to be the best deleterious nsSNPs (Table 1). Some of them were found to be associated with destabilization of EPH-ligand binding domain (R66S, R160H, R104Q), functional impairment of protein kinase domain (G766V, Y779C, N751D, R745Q), alteration of substrate binding sites (P824S), ATP-binding site alteration (G628E, I652M) and substitution of autophosphorylation residues (Y779C, Y596H). So, all these most deleterious nsSNPs predicted by 8 tools (Supplementary Table 3) contain a high probability of disrupting EPHA3 protein functions.

Domain identification of EPHA3 by InterPro

InterPro tool predicted the domain regions of EPHA3 protein. It mainly reported 5 domains: EPH-ligand binding domain (29–207), Fibronectin Type-III domain (325–435), Fibronectin Type-II (436–531), Protein kinase domain (621–882), and SAM domain (911–975). It also predicted the ATP-binding site (627–653) and active site (742–754). Several deleterious nsSNPs were found in the EPH-ligand binding domain (R66S, R160H, R104Q), protein kinase domain (G766V, Y779C, N751D, R745Q), substrate binding sites (P824S), ATP-binding site (G628E, I652M).

Analysis of EPHA3 protein interactions by STRING v11.0:

The server revealed that 10 proteins interact with EPHA3 protein (Fig. 3). These proteins are mainly involved in signaling pathways, protein kinase activity, neuron recognition, and many other pathways. Ephrin type-A ligands, such as EFNA5, EFNA1, EFNA2, EFNA4, EFNA3 ligand peptides bind with EPHA3 receptor protein through glycosyl-phosphatidylinositol (GPI) moiety and Ephrin type B ligands, such as EFNB1, EFNB2, EFNB3, bind to EPHA3 receptor via transmembrane domain. EPHA3 also interacts with the adaptor molecule CRK-II, transforming protein RhoA, and is involved in regulating many vital functions.

The conservational outcome of nsSNPs and phylogenetic analysis of EPHA3 Protein

Highly conserved residues are most likely to involve in protein's structural integrity and functions. So, the conservational profile was evaluated for EPHA3 protein for further analysis. The ConSurf algorithm represented structural and functional conservational levels of all the amino acid residues of EPHA3 protein, illustrated in Fig. 4. ConSurf predicted 77 of 81 nsSNPs to be conserved. The result showed R66, R104, R160, Y779, R842, P846, and G766 as functional residues that are highly conserved and exposed. On the other hand, G628, I652, N751, Y596, D598, R745, and P824 are predicted as highly conserved and buried structural residues. Therefore, these outcomes further confirmed the predicted nsSNPs as high-risk deleterious nsSNPs, which may impact on structural stability and functions of EPHA3 protein. These findings were also verified by multiple sequence analysis (MSA) through the MEGA X package.

Phylogenetic analysis by MEGA X revealed that humans (*Homo sapiens*) and chimpanzees (*Pan troglodytes*) have a high degree of sequence homology, indicating EPHA3 protein is highly conserved in

primates. (Fig. 5)

Functional SNPs prediction in the 3' UTR region by PolymiRTS

PolymiRTS was used to predict nsSNPs in the miRNA target sites as any alteration in these sites may create or disrupt miRNA-mRNA interactions. The server predicted 21 SNPs presented in miRNA target site (Supplementary Table 4), among which 2 SNPs were found as INDELs, whose ancestral alleles are undefined, can alter miRNA target site, and the other 19 SNPs can either disrupt conserved miRNA target site (D) or can create new miRNA target site (C).

Discussion

Several *in-silico* SNPs prediction tools with different algorithms were used in this study concluded multiple nsSNPs in highly conserved regions can potentially have detrimental consequences considering the importance of conserved regions for sustaining protein's structure and function [21]. Eighty-one common nsSNPs predicted by eight different prediction tools are possibly the most targeted mutations for various diseases. Evolutionary conservational profile analysis of all amino acids of EPHA3 was needed to ensure the deleterious impact of selected nsSNPs, and the ConSurf server predicted conservational levels for all amino acids of EPHA3. The server predicted R104Q, R160H, Y779C, R842Q, P846A&T, and G766V functionally exposed nsSNPs as highly conserved residues with a conservancy level of 9 and level 8 for functional conserved residues R66S. On the other hand, it predicted G628E, I652M, N751D, Y596H, D598G, R745Q, P824S nsSNPs as structurally highly conserved buried residues with a conservancy level of 9. As these residues have a high degree of conservancy level, these nsSNPs are more likely to be deleterious in EPHA3 protein and can be used as genetic markers. To further evaluate the ConSurf web server results, multiple sequence alignment and phylogenetic tree analysis were done by the MEGA X package. The analysis showed the sequence homology of 12 different. Phylogenetic tree generation by MEGA X using 1000 bootstraps with a 50% cut-off revealed the close association of *Homo sapiens* and *Pan troglodytes*, indicating catalytic kinase domain of EPHA3 protein is evolutionary conserved mostly in primates. Protein-protein interactions analyzed by STRING v11 revealed that EPHA3 protein is associated with many important functions. Interactions with Ephrin type A and Ephrin type B ligands induce collapse growth cones, adhesion, crucial migration by contact dependant bi-directional pathways. Also, have a pivotal role in forebrain function. Adapter molecule crk-II mediates attachment-induced MAPK8 activation, membrane ruffling, and cell motility in a Rac-dependent manner. The transforming RhoA protein is involved in microtubule-dependent signaling, which is important for the myosin contractile ring formation during cell cycle cytokinesis. The interactions also stimulate PKN2 (serine/threonine-protein kinase N2) kinase activity. It revealed that EPHA3 participates in many important pathways and any disruption may lead to abnormal body functions [22].

In ATP phosphorylation events, Magnesium atoms form chelation with the β and γ phosphate groups of AMP-PNP and with the side chains of the catalytic residues Asn751 (present in a helix) and Asp764 [23,

24]. The catalytic residue Asparagine in the 751st position was found to be substituted with aspartate means a polar neutral amino acid gets substituted with a negatively charged aspartate (D) amino acid residue; this charge difference might hamper protein-protein interactions. So, N751D may contain a high probability to cause hamper in protein functions. EPHA3 activation depends on three events of autophosphorylation [25]; one event occurs in JMS (Juxtamembrane segment) peptide that contains two residues for phosphorylation, JX1 (Tyr596) and JX2(Tyr602), the second event is the phosphorylation of Tyr779 residues, and a third phosphorylation event occurs to the Tyr701 residue. These three phosphorylation events are highly responsible for protein activation. Disruption of these autophosphorylation events causes the progression of various cancers [22].

Tyr596 (Y596) and Tyr779 (Y779) were found to be substituted with Histidine (H) and Cysteine (C), respectively. In the unphosphorylated state of EPHA3, JX2, in concert with JX1(Y596), stabilizes JMS interaction with the core of the kinase domain. Substitution of tyrosine (Y) with positively charged Histidine (H) residue is more likely to hamper autophosphorylation events. Autophosphorylation of Y779, present in the activation loop (AL), involves catalytic activation of EPHA3. The phosphorylated Y779 can bind to SH2 domain-containing Crk adaptor protein, which ultimately initiates a signaling pathway to activate RhoA protein required for growth cone collapse induced by ephrin-A5 by the activation of EphA receptors [26]. Thus mutation in this residue might hamper important biological functions. Attenuation of EPHA3 receptor functions might link to lung cancer [27]. In our studies, it was found that Y779 is substituted with polar cysteine (C) residue, and thus, it may cause inappropriate interactions with surrounding residues, which leads to the deregulation of many important biological events [23]. Therefore, these studies confirmed that the predicted Y596H and Y779C nsSNPs most likely to have deleterious effects on EPHA3 protein.

The highly conserved DFG segment, an extension of the activation loop, is associated with the phosphorylation state of the EPHA3. In the active state, the DFG motif is settled to interact with activating Mg^{2+} ions. From a previous study, it was found that in phosphatase treated samples, which contain well defined JMS segment, activation loop (AL) only order up to DFG⁷⁶⁶ (G⁷⁶⁶->Gly766), and residues Leu767–Ile786 were found to be disordered [28]. This DFG segment is highly associated with the phosphorylation state of the EPHA3 protein. But an SNP was found in the DFG segment, which is in the 766th position. As Glycine (G), a polar hydrophilic residue, is substituted with nonpolar hydrophobic Valine (V) residue, this substitution may hamper catalytic activity, which can lead to various abnormalities [29]. This confirms the predicted G766V nsSNP might have deleterious effects on EPHA3 protein functions.

The sterile alpha motif (SAM) domain, which involves the oligomerization of receptors, interacts with the kinase domain through a linker peptide [30]. This linker peptide is accommodated in a deep pocket formed by α GH loop (residues Tyr841–Pro846) of the kinase domain. Mutations in this domain could affect the dimerization and phosphorylation state of the EPHA3 receptor [31]. In this study, R842Q and P846A, T, R nsSNPs were found to be associated with the pocket. Positively charged arginine (R) residue substituted with polar glutamine (Q) residue, which may cause alteration of charge density and leads to

abnormal interactions. And Proline (P) residue was found to have the probability of being substituted with three different residues, alanine (A), Threonine (T), and arginine (R). These substitutions in this position may disrupt the favorable interactions in the α GH loop and thus linker-kinase domain interaction may not be strong enough to connect SAM domain. As a result, the EPHA3 protein might lose its function to pass the signal, as the oligomerization process is inefficient.

Pretreatment of ATP with protein makes the activation loop (AL) ordered, and this ordered state is stabilized with many H-bonds and Van der Waals interactions. One of the residues, Arg745, is associated with this interaction in stabilizing AL's ordered structure by forming an H-bond with Leu767 [32]. R745Q may be associated with deleterious effects on protein functions as this substitution alter the charge density. The further EPHA3 structure analysis reveals that Asp598 (D598) and Tyr596 (Y596) residues might also be important in the regulatory mechanism as they are involved in ordering the activation loop state. D598G and Y596H nsSNPs were found in this study, and so it suggests that these nsSNPs may have deleterious effects on the activation of EPHA3 protein.

G628 and I652 were found to be present in the ATP-binding domain (627–653) [33]. G628 present in B1-strand of G-loop involved in closed conformation of N-terminal lobe. Phosphorylation events increase the flexibility or dynamics of the N-terminal lobe, which helps move towards the active state of the protein. G628E and I652M nsSNPs may disrupt the ATP binding mechanism and alter EPHA3 protein activation. This could lead to loss of tumor-suppressive effect of EPHA3 protein [22, 27].

R104 in EPH-ligand binding domain associated with a salt bridge with ephrin-A5 and it is one of the most conserved bindings. R104Q, where positively charged arginine (R) is substituted with polar hydrophobic glutamine residue (Q), may ultimately affect receptor-ligand interactions, and thus impaired signaling pathways may most likely occur [34]. R160H and R66S nsSNPs also disrupt the interactions between the EPH-ligand-binding domain and ephrin-A5. Ephrin-A5 and EPHA3 receptor interactions are important for cell communications functions in normal development. Any disruption in these interactions leads to neoplasia [35]. So, R104Q, R160H, and R66S were predicted as deleterious nsSNPs.

The substrate-binding regions are formed by the α F- α G loop/ α G helix [33]. The α F- α G loop (821–829 residues) was found to be associated with an SNP in 824th positions. Nonpolar hydrophobic proline (P) residue, which is highly conserved and buried in the structure, substitutes with polar hydrophilic serine (S) residue. This substitution in concert with neighboring residues may cause alteration of the secondary structure of the substrate-binding domain. Thus, P824S is considered to possess deleterious effects on protein structure and functions.

A small hydrophobic cluster formed by M734, Y810, and F871 surrounding the conserved aspartate in the regulatory spine (known as the R-spine-Asp). Together, these three residues form an interaction network that extends the R-spine to E and I helices in the C-lobe. These PTK-conserved residues are referred to as the 'extended R-spine-network'. Precise positioning of the R-spine-Asp is important for kinase activation and regulation [36]. Thus these 3 residue networks are involved in kinase activation and regulation.

Therefore, Y810H nsSNP may interfere with the activation and regulatory mechanism of EPHA3 protein as polar tyrosine (Y) gets replaced by positively charged histidine residue (H).

Five nsSNPs (N751D, D746E, R750Q, R712C, G784R) were found, from a previously done study [37], in the positions that take part in receptor substrate interaction. So substitutions in these positions may exert deleterious consequences.

Besides, 9 SNPs were found in the fibronectin type 3 domain, among those V415M, L389F, G414V, S327T, R354G, T432N, W345R, R354Q were highly conserved, and 8 SNPs were found in the Fibronectin type 2 domain, among those Y470H and I467K, T was found highly conserved (by both Consurf and MEGAX analysis). So these nsSNPs may alter the fibronectin domain stability and affect EPHA3 protein functions [38]. Previous studies show that G518L mutation was associated with lung adenocarcinoma [39]. So this is plausible that this G518E may be associated with a certain disease as polar hydrophobic glycine (G) residue gets substituted by negatively charged glutamate (E) residue. Differential expression of EPHA3 is associated with many cancers, such as breast cancer, prostate cancer [40], lung cancer, glioblastoma multiforme (GBM) [41], melanoma, sarcoma, renal carcinoma, breast cancer, hematopoietic tumor [42]. EPHA3 is highly expressed in tumor cells and significantly takes part in tumor angiogenesis [43].

EPHA3 can be targeted for the possible treatment of many diseases because EPHA3 is associated with the development and maturation of lymphoid and myeloid cells [44, 45]. As differential expression of the EPHA3 gene lead to many cancers, controlling the expression can be an ideal solution. So, silencing mRNA can be one of the possible routes to control the differential EPHA3 gene expression. siRNA (small interfering RNA) can be designed against a specific conserved region in mRNA, which leads to knocking down the gene [46] a strategy proposed for other diseases as well [47, 48]. Another possible route to limit the effects of abnormal EPHA3 receptors would be designing of antibody against it. ChIIIA4 is one of the antibodies that can work on EPHA3 receptors and inhibit tumor growth and disrupt the function of tumor micro-vessels [43]. KBoo4 is another anti-EPHA3 antibody, which can be used for treating hematologic malignancies [49]. Sitravatinib, Hesperadin, Vandetanib, and Fostamatinib are some inhibitory drugs, collected from the Drug Gene Interaction Database [50], which can be used to limit the detrimental effects of the EPHA3 receptor. Thus, designed siRNA, antibodies, or clinical drugs can be further modified to interact against different regions of EPHA3 receptor where deleterious nsSNPs are located. This will ultimately prevent the abnormal effects of the EPHA3 receptor.

Substitutions in conserved regulatory regions lead to altering the regulations of signaling pathways, which promote abnormal cell differentiation or proliferation. Various damaging nsSNPs were found through this study, N751D, Y596H, Y779C, D598G, P824S, G628E, I652M, R66S, R160H, R104Q were found to have the most deleterious impacts on protein structure and functions as these residues hold very conserved and important positions in the EPHA3 protein and may highly responsible for alternating the expression of EPHA3 gene, which can cause cancers or tumor growth.

Conclusion

This study unravels several nsSNPs that can be deleterious and have mutational impacts on protein structure and functions of the EPHA3 protein using an extensive computational approach. 77 nsSNPs showed a high level of conservancy level, and thus, they can be used as genetic markers that will facilitate the finding of diseases associated with EPHA3 protein. This study screened out significant mutational alterations which may alter conformational changes of the EPHA3 protein structure which ultimately lead to disruption of protein functions. Therefore, this comprehensive study will ultimately open many possibilities to work on diseases or cancers associated with EPHA3 protein and can provide necessary supports in the development of target-specific therapeutic agents or drugs.

Declarations

Declaration of Competing Interest

The authors declare that they have no competing interests.

Ethical approval

This study doesn't involve any human or animal-related experiment.

Authors Contribution

Dipankor Chatterjee: Writing, Analysis

Umar Faruq Chowdhury: Writing, Analysis

Mohammad Umer Sharif Shohan: Writing, Conceptualization

Md Mohasin: Review

Yearul Kabir: Review, Supervision.

Funding

The research did not receive any funds.

References

1. O'Leary DD, Wilkinson DG (1999) Eph receptors and ephrins in neural development. *Curr Opin Neurobiol* 9:65–73. [https://doi.org/10.1016/S0959-4388\(99\)80008-7](https://doi.org/10.1016/S0959-4388(99)80008-7)
2. Janes PW (2017) EPHA3 (EPH receptor A3). *Atlas Genet Cytogenet Oncol Haematol* 3:264–272. <https://doi.org/10.4267/2042/62775>
3. Lawrenson ID, Wimmer-Kleikamp SH, Lock P et al (2002) Ephrin-A5 induces rounding, blebbing and de-adhesion of EphA3-expressing 293T and melanoma cells by CrkII and Rho-mediated signalling. *J Cell Sci* 115:1059 LP – 1072

4. Hussain MRM, Shaik NA, Al-Aama JY et al (2012) In silico analysis of Single Nucleotide Polymorphisms (SNPs) in human BRAF gene. *Gene* 508:188–196.
<https://doi.org/10.1016/j.gene.2012.07.014>
5. De Alencar SA, Lopes JCD (2010) A Comprehensive in silico analysis of the functional and structural impact of SNPs in the IGF1R gene. *J Biomed Biotechnol* 2010:.
<https://doi.org/10.1155/2010/715139>
6. Sherry ST, Ward MH, Kholodov M et al (2001) DbSNP: The NCBI database of genetic variation. *Nucleic Acids Res* 29:308–311. <https://doi.org/10.1093/nar/29.1.308>
7. Ng PC, Henikoff S (2003) SIFT: Predicting amino acid changes that affect protein function. *Nucleic Acids Res* 31:3812–3814. <https://doi.org/10.1093/nar/gkg509>
8. Adzhubei IA, Schmidt S, Peshkin L et al (2010) A method and server for predicting damaging missense mutations. *Nat Methods* 7:248–249
9. Capriotti E, Fariselli P, Calabrese R, Casadio R (2005) Predicting protein stability changes from sequences using support vector machines. *Bioinformatics* 21:.
<https://doi.org/10.1093/bioinformatics/bti1109>
10. Choi Y, Sims GE, Murphy S et al (2012) Predicting the Functional Effect of Amino Acid Substitutions and Indels. *PLoS One* 7:.
<https://doi.org/10.1371/journal.pone.0046688>
11. Calabrese R, Capriotti E, Fariselli P et al (2009) Functional annotations improve the predictive score of human disease-related mutations in proteins. *Hum Mutat* 30:1237–1244.
<https://doi.org/10.1002/humu.21047>
12. Bromberg Y, Yachdav G, Rost B (2008) SNAP predicts effect of mutations on protein function. *Bioinformatics* 24:2397–2398. <https://doi.org/10.1093/bioinformatics/btn435>
13. Niroula A, Urolagin S, Vihinen M (2015) PON-P2: Prediction Method for Fast and Reliable Identification of Harmful Variants. *PLoS One* 10:e0117380.
<https://doi.org/10.1371/journal.pone.0117380>
14. Reva B, Antipin Y, Sander C (2011) Predicting the functional impact of protein mutations: Application to cancer genomics. *Nucleic Acids Res* 39:e118–e118. <https://doi.org/10.1093/nar/gkr407>
15. Hunter S, Apweiler R, Attwood TK et al (2009) InterPro: The integrative protein signature database. *Nucleic Acids Res* 37:211–215. <https://doi.org/10.1093/nar/gkn785>
16. Szklarczyk D, Gable AL, Lyon D et al (2019) STRING v11: protein–protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res* 47:D607–D613. <https://doi.org/10.1093/nar/gky1131>
17. Ashkenazy H, Erez E, Martz E et al (2010) ConSurf 2010: calculating evolutionary conservation in sequence and structure of proteins and nucleic acids. *Nucleic Acids Res* 38:W529–W533.
<https://doi.org/10.1093/nar/gkq399>
18. Pupko T, Bell RE, Mayrose I et al (2002) Rate4Site: An algorithmic tool for the identification of functional regions in proteins by surface mapping of evolutionary determinants within their homologues. In: *Bioinformatics*. Oxford University Press

19. Kumar S, Stecher G, Li M et al (2018) MEGA X: Molecular evolutionary genetics analysis across computing platforms. *Mol Biol Evol* 35:1547–1549. <https://doi.org/10.1093/molbev/msy096>
20. Pickrell JK, Marioni JC, Pai AA et al (2010) Understanding mechanisms underlying human gene expression variation with RNA sequencing. *Nature* 464:768–772. <https://doi.org/10.1038/nature08872>
21. Greene LH, Chrysina ED, Irons LI et al (2009) Role of conserved residues in structure and stability: Tryptophans of human serum retinol-binding protein, a model for the lipocalin superfamily. *Protein Sci* 10:2301–2316. <https://doi.org/10.1110/ps.22901>
22. Lisabeth EM, Fernandez C, Pasquale EB (2012) Cancer somatic mutations disrupt functions of the EphA3 receptor tyrosine kinase through multiple mechanisms. *Biochemistry* 51:1464–1475. <https://doi.org/10.1021/bi2014079>
23. Davis TL, Walker JR, Loppnau P et al (2008) Autoregulation by the Juxtamembrane Region of the Human Ephrin Receptor Tyrosine Kinase A3 (EphA3). *Structure* 16:873–884. <https://doi.org/10.1016/j.str.2008.03.008>
24. Qian KC, Wang L, Hickey ER et al (2005) Structural basis of constitutive activity and a unique nucleotide binding mode of human Pim-1 kinase. *J Biol Chem* 280:6130–6137. <https://doi.org/10.1074/jbc.M409123200>
25. Binns KL, Taylor PP, Sicheri F et al (2000) Phosphorylation of Tyrosine Residues in the Kinase Domain and Juxtamembrane Region Regulates the Biological and Catalytic Activities of Eph Receptors. *Mol Cell Biol* 20:4791–4805. <https://doi.org/10.1128/mcb.20.13.4791-4805.2000>
26. Singh A, Winterbottom E, Daar IO (2012) Eph/ephrin signaling in cell-cell and cell-substrate adhesion. *Front Biosci* 17:473–497. <https://doi.org/10.2741/3939>
27. Zhuang G, Song W, Amato K et al (2012) Effects of cancer-associated EPHA3 mutations on lung cancer. *J Natl Cancer Inst* 104:1182–1197. <https://doi.org/10.1093/jnci/djs297>
28. Taylor SS, Kornev AP (2011) Protein kinases: evolution of dynamic regulatory proteins. *Trends Biochem Sci* 36:65–77. <https://doi.org/10.1016/j.tibs.2010.09.006>
29. Valsesia A, Rimoldi D, Martinet D et al (2011) Network-Guided Analysis of Genes with Altered Somatic Copy Number and Gene Expression Reveals Pathways Commonly Perturbed in Metastatic Melanoma. *PLoS One* 6:e18369. <https://doi.org/10.1371/journal.pone.0018369>
30. Stapleton D, Balan I, Pawson T, Sicheri F (1999) The crystal structure of an Eph receptor SAM domain reveals a mechanism for modular dimerization. *Nat Struct Biol* 6:44–49. <https://doi.org/10.1038/4917>
31. Singh DR, Cao QQ, King C et al (2015) Unliganded EphA3 dimerization promoted by the SAM domain. *Biochem J* 471:101–109. <https://doi.org/10.1042/BJ20150433>
32. Wiesner S, Wybenga-Groot LE, Warner N et al (2006) A change in conformational dynamics underlies the activation of Eph receptor tyrosine kinases. *EMBO J* 25:4686–4696. <https://doi.org/10.1038/sj.emboj.7601315>

33. Kwon A, John M, Ruan Z, Kannan N (2018) Coupled regulation by the juxtamembrane and sterile α motif (SAM) linker is a hallmark of ephrin tyrosine kinase evolution. *J Biol Chem* 293:5102–5116. <https://doi.org/10.1074/jbc.RA117.001296>
34. Kou CTJ, Kandpal RP (2018) Differential Expression Patterns of Eph Receptors and Ephrin Ligands in Human Cancers. *Biomed Res. Int.* 2018
35. Smith FM, Vearing C, Lackmann M et al (2004) Dissecting the EphA3/Ephrin-A5 Interactions Using a Novel Functional Mutagenesis Screen. *J Biol Chem* 279:9522–9531. <https://doi.org/10.1074/jbc.M309326200>
36. Mohanty S, Oruganty K, Kwon A et al (2016) Hydrophobic Core Variations Provide a Structural Framework for Tyrosine Kinase Evolution and Functional Specialization. *PLoS Genet* 12:. <https://doi.org/10.1371/journal.pgen.1005885>
37. Davis TL, Walker JR, Allali-Hassani A et al (2009) Structural recognition of an optimized substrate for the ephrin family of receptor tyrosine kinases. *FEBS J* 276:4395–4404. <https://doi.org/10.1111/j.1742-4658.2009.07147.x>
38. Balakrishnan A, Bleeker FE, Lamba S et al (2007) Novel somatic and germline mutations in cancer candidate genes in glioblastoma, melanoma, and pancreatic carcinoma. *Cancer Res* 67:3545–3550. <https://doi.org/10.1158/0008-5472.CAN-07-0065>
39. Falivelli G, Lisabeth EM, De La Torre ER et al (2013) Attenuation of Eph receptor kinase activation in cancer cells by coexpressed ephrin ligands. *PLoS One* 8:. <https://doi.org/10.1371/journal.pone.0081445>
40. Singh AP, Bafna S, Chaudhary K et al (2008) Genome-wide expression profiling reveals transcriptomic variation and perturbed gene networks in androgen-dependent and androgen-independent prostate cancer cells. *Cancer Lett* 259:28–38. <https://doi.org/10.1016/j.canlet.2007.09.018>
41. Day BW, Stringer BW, Boyd AW (2014) Eph receptors as therapeutic targets in glioblastoma. *Br J Cancer* 111:1255–1261
42. Dottori M, Down M, Hüttmann A et al (1999) Cloning and characterization of EphA3 (Hek) gene promoter: DNA methylation regulates expression in hematopoietic tumor cells. *Blood* 94:2477–2486. https://doi.org/10.1182/blood.v94.7.2477.419k13_2477_2486
43. Vail ME, Murone C, Tan A et al (2014) Targeting EphA3 Inhibits Cancer Growth by Disrupting the Tumor Stromal Microenvironment. *Cancer Res* 74:4470 LP – 4481. <https://doi.org/10.1158/0008-5472.CAN-14-0218>
44. Mäkinen T, Adams RH, Bailey J et al (2005) PDZ interaction site in ephrinB2 is required for the remodeling of lymphatic vasculature. *Genes Dev* 19:397–410. <https://doi.org/10.1101/gad.330105>
45. Adams RH, Klein R (2000) Eph receptors and Ephrin ligands: Essential mediators of vascular development. *Trends Cardiovasc Med* 10:183–188
46. Caivano A, La Rocca F, Laurenzana I et al (2017) EphA3 acts as proangiogenic factor in multiple myeloma. *Oncotarget* 8:34298–34309. <https://doi.org/10.18632/oncotarget.16100>

47. Uludağ H, Parent K, Aliabadi HM, Haddadi A (2020) Prospects for RNAi Therapy of COVID-19. *Front. Bioeng Biotechnol* 8:916
48. Chowdhury UF, Sharif Shohan MU, Hoque KI et al (2021) A computational approach to design potential siRNA molecules as a prospective tool for silencing nucleocapsid phosphoprotein and surface glycoprotein gene of SARS-CoV-2. *Genomics* 113:331–343. <https://doi.org/10.1016/j.ygeno.2020.12.021>
49. Swords RT, Greenberg PL, Wei AH et al (2016) KB004, a first in class monoclonal antibody targeting the receptor tyrosine kinase EphA3, in patients with advanced hematologic malignancies: Results from a phase 1 study. *Leuk Res* 50:123–131. <https://doi.org/10.1016/j.leukres.2016.09.012>
50. Cotto KC, Wagner AH, Feng YY et al (2018) DGIdb 3.0: A redesign and expansion of the drug-gene interaction database. *Nucleic Acids Res* 46:D1068–D1073. <https://doi.org/10.1093/nar/gkx1143>

Tables

Table 1. Prediction of most deleterious nsSNPs by various tools. Abbreviations: PD (Probably Damaging), Del (Deleterious).

nsSNP Id.	Amino Acid Substitution	SIFT	PolyPhen-2	I-Mutant 3.0	PROVEAN	SNPs&GO	SNAP-2	PON-P2	Mutant Assessor
rs1486190561	R312Q	Damaging	PD	Large Decrease	Del	Disease	Effect	Pathogenic	Medium
rs1472882999	Y596H	Damaging	PD	Large Decrease	Del	Disease	Effect	Pathogenic	High
rs1469427775	D149G	Damaging	PD	Large Decrease	Del	Disease	Effect	Pathogenic	Medium
rs1464031640	G955V	Damaging	PD	Large Decrease	Del	Disease	Effect	Pathogenic	Medium
rs1439127416	P605S	Damaging	PD	Large Decrease	Del	Disease	Effect	Unknown	High
rs1432350325	V415M	Damaging	PD	Large Decrease	Del	Disease	Effect	Pathogenic	High
rs1426641543	P885T	Damaging	PD	Large Decrease	Del	Disease	Effect	Pathogenic	Medium
	P885S		PD		Del	Disease	Effect	Pathogenic	Medium
rs1424411876	P869S	Damaging	PD	Large Decrease	Del	Disease	Effect	Pathogenic	Medium
rs1400852070	L389F	Damaging	PD	Large Decrease	Del	Disease	Effect	Pathogenic	Medium
rs1400426073	T67N	Damaging	PD	Large Decrease	Del	Disease	Effect	Pathogenic	Medium
rs1391704321	Y470H	Damaging	PD	Large Decrease	Del	Disease	Effect	Pathogenic	High
rs1383248514	A748T	Damaging	PD	Large Decrease	Del	Disease	Effect	Pathogenic	Medium
rs1380975088	R868S	Damaging	PD	Large Decrease	Del	Disease	Effect	Pathogenic	High
rs1378343025	V197G	Damaging	PD	Large Decrease	Del	Disease	Effect	Unknown	Medium
rs1366099808	G766V	Damaging	PD	Large Decrease	Del	Disease	Effect	Pathogenic	High
rs1362369938	G518E	Damaging	PD	Large Decrease	Del	Disease	Effect	Pathogenic	High
rs1354537650	R66S	Damaging	PD	Large Decrease	Del	Disease	Effect	Pathogenic	Medium
rs1347519802	N107H	Damaging	PD	Large Decrease	Del	Disease	Effect	Pathogenic	Medium
rs1334127938	G414V	Damaging	PD	Large Decrease	Del	Disease	Effect	Pathogenic	High
rs1315586801	S327T	Damaging	PD	Large Decrease	Del	Disease	Effect	Pathogenic	Medium
rs1315283374	P846R	Damaging	PD	Large Decrease	Del	Disease	Effect	Pathogenic	Medium
rs1310443133	G628E	Damaging	PD	Large Decrease	Del	Disease	Effect	Pathogenic	High
rs1294076209	K100N	Damaging	PD	Large Decrease	Del	Disease	Effect	Unknown	Medium
rs1288096834	D746E	Damaging	PD	Large Decrease	Del	Disease	Effect	Unknown	High
rs1283969718	Y810H	Damaging	PD	Large Decrease	Del	Disease	Effect	Pathogenic	Medium
rs1279990429	Y68H	Damaging	PD	Large	Del	Disease	Effect	Pathogenic	Medium

				Decrease					
rs1270443958	A182T	Damaging	PD	Large Decrease	Del	Disease	Effect	Pathogenic	Medium
rs1262606547	I652M	Damaging	PD	Large Decrease	Del	Disease	Effect	Pathogenic	Medium
rs1242408900	I874T	Damaging	PD	Large Decrease	Del	Disease	Effect	Pathogenic	High
rs1230871085	A148G	Damaging	PD	Large Decrease	Del	Disease	Effect	Unknown	Medium
rs1229614068	V252A	Damaging	PD	Large Decrease	Del	Disease	Effect	Pathogenic	High
rs1203973847	R799C	Damaging	PD	Large Decrease	Del	Disease	Effect	Pathogenic	Medium
rs1190951059	G228R	Damaging	PD	Large Decrease	Del	Disease	Effect	Pathogenic	High

Table 1. Prediction of most deleterious nsSNPs by various tools. Abbreviations: PD (Probably Damaging), Del (Deleterious). (Continued)

nsSNP Id.	Amino Acid Substitution	SIFT	PolyPhen-2	I-Mutant 3.0	PROVEAN	SNPs&GO	SNAP-2	PON-P2	Mutant Assessor
rs1164159108	L164F	Damaging	PD	Large Decrease	Del	Disease	Effect	Unknown	Medium
rs998163895	W919C	Damaging	PD	Large Decrease	Del	Disease	Effect	Pathogenic	Medium
rs995238704	R381C	Damaging	PD	Large Decrease	Del	Disease	Effect	Unknown	Medium
rs866959883	G784E	Damaging	PD	Large Decrease	Del	Disease	Effect	Pathogenic	Medium
rs866821420	D156N	Damaging	PD	Large Decrease	Del	Disease	Effect	Pathogenic	Medium
rs866025221	E615K	Damaging	PD	Large Decrease	Del	Disease	Effect	Pathogenic	High
rs780417089	R750Q	Damaging	PD	Large Decrease	Del	Disease	Effect	Pathogenic	Medium
rs779599137	R354G	Damaging	PD	Large Decrease	Del	Disease	Effect	Pathogenic	High
rs778067909	W919L	Damaging	PD	Large Decrease	Del	Disease	Effect	Pathogenic	Medium
rs777211567	R712C	Damaging	PD	Large Decrease	Del	Disease	Effect	Pathogenic	Medium
rs775011935	P463R	Damaging	PD	Large Decrease	Del	Disease	Effect	Pathogenic	Medium
rs774002419	F667S	Damaging	PD	Large Decrease	Del	Disease	Effect	Pathogenic	Medium
rs773980261	Y779C	Damaging	PD	Large Decrease	Del	Disease	Effect	Pathogenic	Medium
rs769554825	V195A	Damaging	PD	Large Decrease	Del	Disease	Effect	Pathogenic	Medium
rs768163795	R160C	Damaging	PD	Large Decrease	Del	Disease	Effect	Pathogenic	Medium
rs766046211	T432N	Damaging	PD	Large Decrease	Del	Disease	Effect	Pathogenic	High
rs765969815	S805N	Damaging	PD	Large Decrease	Del	Disease	Effect	Pathogenic	Medium
rs763618226	D598G	Damaging	PD	Large Decrease	Del	Disease	Effect	Pathogenic	High
rs761576191	P253H	Damaging	PD	Large Decrease	Del	Disease	Effect	Pathogenic	High
rs758438939	L499F	Damaging	PD	Large Decrease	Del	Disease	Effect	Pathogenic	High
rs758183213	R842Q	Damaging	PD	Large Decrease	Del	Disease	Effect	Pathogenic	Medium
rs757274748	P824S	Damaging	PD	Large Decrease	Del	Disease	Effect	Pathogenic	Medium
rs756739935	R745Q	Damaging	PD	Large Decrease	Del	Disease	Effect	Pathogenic	High
rs756018331	L42P	Damaging	PD	Large	Del	Disease	Effect	Pathogenic	Medium

				Decrease					
rs755409303	I754N	Damaging	PD	Large Decrease	Del	Disease	Effect	Pathogenic	Medium
rs754635213	G705A	Damaging	PD	Large Decrease	Del	Disease	Effect	Pathogenic	High
rs753700696	R104Q	Damaging	PD	Large Decrease	Del	Disease	Effect	Pathogenic	Medium
rs752560194	R883G	Damaging	PD	Large Decrease	Del	Disease	Effect	Pathogenic	Medium
rs751094345	E150K	Damaging	PD	Large Decrease	Del	Disease	Effect	Pathogenic	Medium
rs750832814	L468S	Damaging	PD	Large Decrease	Del	Disease	Effect	Pathogenic	Medium
rs750779844	R160H	Damaging	PD	Large Decrease	Del	Disease	Effect	Pathogenic	Medium
rs749646074	N751D	Damaging	PD	Large Decrease	Del	Disease	Effect	Pathogenic	High
rs748606864	L668P	Damaging	PD	Large Decrease	Del	Disease	Effect	Pathogenic	Medium

Table 1. Prediction of most deleterious nsSNPs by various tools. Abbreviations: PD (Probably Damaging), Del (Deleterious). (Continued)

nsSNP Id.	Amino Acid Substitution	SIFT	PolyPhen-2	I-Mutant 3.0	PROVEAN	SNPs&GO	SNAP-2	PON-P2	Mutant Assessor
rs570604469	I467K	Damaging	PD	Large	Del	Disease	Effect	Pathogenic	High
	I467T		PD	Decrease	Del	Disease	Effect	Pathogenic	High
rs556754042	G840V	Damaging	PD	Large Decrease	Del	Disease	Effect	Pathogenic	Medium
rs374701071	W345R	Damaging	PD	Large Decrease	Del	Disease	Effect	Pathogenic	High
rs374211029	E41K	Damaging	PD	Large Decrease	Del	Disease	Effect	Unknown	Medium
rs370552845	Y474C	Damaging	PD	Large Decrease	Del	Disease	Effect	Pathogenic	Medium
rs202143347	Y124C	Damaging	PD	Large Decrease	Del	Disease	Effect	Pathogenic	Medium
rs200071736	L685P	Damaging	PD	Large Decrease	Del	Disease	Effect	Pathogenic	Medium
rs144727411	R354Q	Damaging	PD	Large Decrease	Del	Disease	Effect	Pathogenic	High
rs142030914	P846A	Damaging	PD	Large	Del	Disease	Effect	Unknown	Medium
	P846T		PD	Decrease	Del		Effect	Pathogenic	Medium
rs141287594	P173L	Damaging	PD	Large Decrease	Del	Disease	Effect	Pathogenic	Medium
rs138376716	E265G	Damaging	PD	Large Decrease	Del	Disease	Effect	Pathogenic	Medium
rs73139269	F801S	Damaging	PD	Large Decrease	Del	Disease	Effect	Pathogenic	Medium

Figures

Figure 1

Designed workflow of methodology

Figure 2
Percentile distribution of EPHA3 gene SNPs

Figure 3

Exploration of protein-protein interactions through STRING v11 server

Figure 5

Analysis of the phylogenetic tree of Human EPHA3 protein.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementaryTable1.xlsx](#)
- [SupplementaryTable2.xlsx](#)
- [SupplementaryTable3.xlsx](#)
- [SupplementaryTable4.xlsx](#)