BIOINFORMATICS EVALUATION OF *CRISP2* GENE SNPs AND THEIR IMPACTS ON PROTEIN

Zaid A. Hussein **Abdulkareem A. Al-Kazaz**
Researcher **Prof.**
Dept Biotechn. Coll. Scie. University of Baghdad. Baghdad, Iraq
Corresponding author E-mail: zaid.a@sc.uobaghdad.edu.iq

ABSTRACT

This study evaluated the *CRISP2* gene's functional single nucleotide polymorphisms, and its results may be advantageous for future population-based studies and early diagnostic discoveries, particularly in developing effective treatments. The *CRISP2* gene encodes a secretory protein with a high cysteine content, which belongs to the family of cysteine-rich secretory proteins (CRISP). SNPs are genetic variations that may affect a protein's structure or functionality. Prior to carrying out a broader population investigation, it is possible to evaluate suspected functional SNPs since it is challenging to uncover functional SNPs in disease-linked genes. As a result, using various bioinformatic prediction models, the potentially harmful three SNPs of the *CRISP2* gene were predicted in this in-silico study from the neutral ones. Out of a pool of 260 nsSNPs, three SNPs (L56V, M176I, and C196R) been selected to anticipate their impacts on functions and structures along with their capability to impair protein stability. Actually, two of the three SNPs in the *CRISP2* gene L56V and C196R were identified as possibly detrimental, although M176I was not. But all of these SNPs dropped significantly protein stability, per the I-Mutant suite.

Keywords: Polymorphism, PROVEN, Polyphen-2, PHD-SNP, 1-Mutant suite.

تقييم معلوماتي– حياتي لتعدد اشكال النيوكليوتيدة المفردة لجين *CRISP2* وتأثيراتها على البروتين

زيد علي حسين عبد الكريم عبد الرازق القزاز
باحث أستاذ
قسم التقنيات الاحيائية / كلية العلوم / جامعة بغداد

المستخلص

تهدف الدراسة الى تقييم تعدد الأشكال الوظيفية للمورث *CRISP2* ، النتائج ستكون بالتأكيد مفيدة للآفاق المستقبلية المتعلقة بالدراسات السكانية الكبيرة. بالإضافة إلى الاكتشافات التشخيصية المبكرة، لا سيما في إكتشاف العلاج المناسب. يشفر البروتين الإفرازي الغني بالسيستين بواسطة المورث *CRISP2*. ينتمي هذا البروتين إلى عائلة البروتين الإفرازية الغنية بالسيستين .قد تؤدي الاختلافات الجينية من خلال تعدد اشكال النيوكليوتيدة المفردة إلى تأثيرات في بنية و / أو وظيفة البروتين. نظرًا لأن توصيف تعدد الأشكال الوظيفية في المورثات المرتبطة بالأمراض يمثل تحديًا كبيرًا، فمن الممكن تحليل تعدد الأشكال الوظيفية المفترضة قبل إجراء دراسات سكانية أكبر. نتيجة لذلك، في هذه الدراسة في السيليكو، توقعنا الضرر المحتمل لثلاثة طفرات في المورث *CRISP2* من المورثات المحايدة باستخدام أدوات التنبؤ المعلوماتية الحيوية المختلفة. من إجمالي 260 طفرة، تم اختيار ثلاثة طفرات (M176I, L56V و C196R) من أصل 260 للتنبؤ بآثارها الضارة على الوظائف / البنية والقدرة على تقليل استقرار البروتين. في الواقع، اثنان من الاشكال (L56V و C196R) من المورث *CRISP2* أظهرتا امكانية الضرر، ب في حين أن M176I لم تكن كذلك. وأدت جميع الأشكال إلى انخفاض استقرار البروتين، وفقًا لبرنامج I-Mutant suite.

كلمات مفتاحية: تعدد الأشكال, PROVEN, Polyphen-2, PHD-SNP, 1-Mutant suite.

Received:11/7/2021, Accepted:13/10/2021

INTRODUCTION

All societies are affected by the worldwide public health issue of infertility, which is one of the most common disorders among individuals aged between 20 and 45 (21); and has an influence on their families and communities. Around 48.5 million (15–20%) of reproductive-age couples worldwide experience infertility, based on the World Health Organization (40). in a society where trauma and stress were triggered by psychological, economical, and medical factors, especially in one like ours where childbearing is highly valued (31). A global decline in the quality of young, healthy men's sperm is being supported by increasing data (45). Environmental rather than inherited causes are more likely to be responsible for this dramatic decline (44, 35). In approximately 15% of cases of male infertility, genetic factors such as single nucleotide polymorphisms (SNPs), chromosomal abnormalities, and Y chromosome microdeletions could be found (24). The set of genes linked to male infertility has grown (42). As a result of advances in genomics, our knowledge of how genes impact male health and fertility has improved (33). The *CRISP2* gene, which has been identified as a protein-coding gene on chromosome 6p21.3, is the subject of this study. It is over 21 KB in size and has ten coding exons. A 243-amino acid protein is encoded by exons 4–10 (18). Known as cysteine-rich secretory protein 2 (CRISP2), it belongs to the Cysteine Rich Secretory Proteins (CRISPs), a subgroup of the CAP superfamily that is distributed widely among vertebrates and exhibits high levels of conservation (23). Since it influences the morphology and motility of male ejaculated spermatozoa as the only CRISP subfamily member expressed in the mammalian testis, the CRISP2 protein has received interest. The CRISP2 protein is not glycosylated, and androgens do not govern its expression (14). The sperm's progressive motility and acrosome reaction are regulated by the CRISP2 protein. further linked to the CatSper subunit, which generates optimal sperm flagellar beating and regulates ryanodine receptors (25). Single nucleotide polymorphisms (SNPs) account for 90% of human genetic variability (4). By influencing proteins, messenger RNA shape (stability), and promoter activity (gene expression), they might contribute to disease (5). SNPs are the most often employed molecular markers in genetic disease research (30), due to their extensive distribution throughout a given genome and their low cost when compared to others (12). So, detecting and evaluating different gene variants could lead to a deeper explanation of their impact on gene function and physical wellbeing (32). Several computational-based approaches for identifying potential and crucially significant variants prior to testing in vitro or in vivo settings have been developed to lessen this enormous effort. In this situation, applying specific algorithms, the in-silico approach is a practical tool for sorting hazardous from benign SNPs. An array of databases can be used to investigate the entire impact of polymorphism, including both functional and structural changes. In order to predict the detrimental influences of three variants in the *CRISP2* gene and related consequences on protein structure and function, this work has made use of publicly available datasets and freely downloadable bioinformatics tools. This preliminary study of these *CRISP2* gene SNPs will be helpful in subsequent studies to cure infertility.

MATERIALS AND METHODS

Three SNPs from the CRISP2 gene were selected for this study: leucine 59 valine (rs1765509750) in exon 5, methionine 176 isoleucine (rs533319863), and cysteine 196 arginine (rs36069724) in exon 9. Among the computational tools employed were I-Mutant Suite, Poly-Phen-2, PhD-SNP, and PROVEAN.

Retrieving of SNPs

Data about the human *CRISP2* gene and its protein sequence were obtained via NCBI (https://www.ncbi.nlm.nih.gov/) (FASTA format). For computational analysis, details regarding the *CRISP2* SNPs of concern (SNP ID, protein accession number, location, residue change, and global minor allele frequency (MAF)) were gathered from the NCBI dbSNP database, which is the most extensive SNP database (http://www.ncbi.nlm.nih.gov/SNP/) (7).

Identification and prediction of the impact of detrimental SNPs: PROVEAN, PolyPhen-2, and PhD-SNP were used successively to evaluate the functional and structural effects of the three *CRISP2* gene SNPs

PROVEAN

Protein Variation Effect Analyzer (http://provean.jcvi.org/index.php) is a software program that predicts if a single amino acid change, insertion, or deletion has a detrimental or neutral effect on a protein's biological function (10). It employs a scoring strategy based on alignment. If the PROVEAN score is ≤ -2.5, the amino acid variant has a detrimental impact on the protein, whereas variants with scores > -2.5 are deemed to have a neutral effect on the protein (21).

PolyPhen-2

Polyphen-2-Polymorphism Phenotyping v2 (http://genetics.bwh.harvard.edu/pph2/) is an online bioinformatics software that predicts the consequence of an amino acid substitution on a protein's structure and, hence, function (2). This prediction is made on the sequence in addition to the way replacement alters structure and phylogeny. The multiple sequence alignment of a 3D protein structure forms the basis for this program's methodology. It incorporates information from numerous protein structural databases. Following that, the position-specific independent count (PSIC) score for each variant is calculated. As an input query for PolyPhen-2, the server got a protein sequence, a database ID/access number, and data on amino acid substitutions. The Poly-Phen-2 score, it ranges from 0.0 to 1.0 and reflects whether a particular amino acid replacement is tolerable or damaging, quantifies the likelihood that a substitution is detrimental. The outcomes are categorized as benign (0–0.15), potentially damaging (0.15-0.85), and probably damaging (0.85–1) (1).

PHD‑SNP

The website (http://snps.biofold.org/phd-snp/phd-snp.html) contains a tool called the Predictor of Human Detrimental Single Nucleotide Polymorphisms. It determines if the new phenotype obtained from an SNP is a neutral polymorphism or connected to genetic illnesses in people using support vector machines (SVMs), which are the basis for the method. Input queries for the protein sequence, the mutation's location, and the mutant residue were all required (8).

Prediction of nsSNPs' effects on protein stability by the I-Mutant Suite: The influence of Variants on the free energy change value (Delta Delta G: DDG) depending on the protein's sequence or tertiary structure was simulated using the I-Mutant Suite (http://gpcr2.biocomp.unibo.it/cgi/predictors/I-Mutant3.0/I-Mutant3.0.cgi) (43) By comparing the free energies of folded mutant and wild-type structures, it is possible to predict how the Gibbs free energy, or Delta Delta G (DDG), would change (20). I-Mutant, which aids in foretelling any change in protein stability after it has been altered, computes the sign of protein stability changes and correlates Delta-Delta G values via support vector machines (SVMs)-based web servers. The software also gives data from ProTherm, a huge library of experimental data on protein mutations. The outcomes are predicted with a reliability index (RI) ranging from 0 to 10, with 10 being the most reliable (6). The tool's output is a DDG value that is categorized as neutral ($(-0.5 \leq$ DDG ≤ 0.5 kcal/mol), largely stable (DDG $>$ 0.5 kcal/mol), or unstable (DDG < -0.5 kcal/mol) (9). In order to determine how three SNPs (L59V, M176I, and C196R) might alter protein stability, the CRISP2 protein sequence was submitted to the I-Mutant software. Conditions of 25 °C and pH 7.0 were applied to all submissions.

RESULTS AND DISCUSSION

Understanding the disease's molecular origins at a molecular scale demands a considerable amount of research and time. There is a necessity for in silico research that can more rapidly and reliably analyze functional SNPs owing to the drawbacks of conventional methodologies (3, 37). A deep understanding of genetic variations in disease susceptibility and therapeutic response may well be achieved by the use of in silico algorithms to predict the phenotypic consequences of nsSNPs. The output is the sorting of deleterious SNPs from benign ones, despite differences in the input and output of these methods as well prediction limitations. The total predictive potential of these methods increases if they are combined. However, the validation of these prediction

algorithms requires supporting data in light of experimental research (22, 34, 41). The investigation of *CRISP2* polymorphisms in fertile males and patients with asthenozoospermia in the Iraqi population is being done in silico for the first time. For this prediction study, we looked at three CRISP2 mutants (L56V/rs1765509750, M176I/rs533319863, and C196R/rs36069724). We employed three distinct in silico methods, including PROVEAN, Poly-Phen-2, and PhD-SNP, to identify the detrimental nature of the structure and/or function of the CRISP2 protein that may be contributing to fertility issues in the variations. Protein stability was additionally predicted using I-Mutant 2.0.

Retrieval of nsSNPs from the NCBI SNP database: The dbSNP database was used to find the variants for the human *CRISP2* gene (dbSNP NCBI: https://www.ncbi.nlm.nih.gov/snp/?term=CRISP2) (7). A total of 15846 SNPs were detected in the human *CRISP2* gene, of which 260 were missense (nsSNP), accounting for just 1.64% of all SNPs known. In addition, there were 757 non-coding transcripts, 92 synonymous, 14471 intronic, 1 initiator codon variant, 1 inframe insertion, 1 inframe indel, and 4 inframe deletions (Fig. 1). Only three CRISP2 nsSNPs were selected for this investigation (L59V, M176I, and C196R).

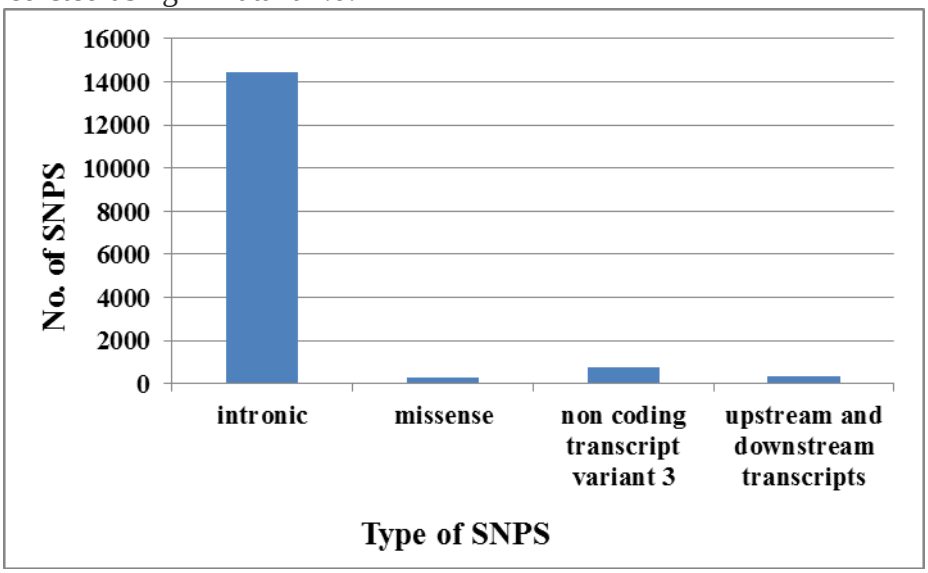

Figure 1. Distribution of *CRISP2* gene SNPs for various functional groups was gathered from the dbSNP database

Prediction of deleterious effects of SNPs on the structure and/or function of CRISP2 protein The majority of human non-synonymous single nucleotide polymorphisms (nsSNPs) exhibit both genetic and phenotypic differences. The nsSNP study's objective is to clarify the genetic link among a range of difficult human disorders and these variants (29). The current computational study was carried out in order to identify the single amino acid variants that are responsible for the changes in the CRISP2 protein's functional and structural properties. To detect the most harmful mutations, three missense variants in the protein-coding region of *CRISP2* were taken from the NCBI dbSNP database and submitted to different functional alteration prediction models, including PROVEAN, PolyPhen 2.0, and PhDSNP.

PROVEN

PROVEAN is a tool for identifying functionally significant nsSNPs. PROVEAN categorized L59V and C196R nsSNPs as deleterious when the threshold value was less than -2.5 (-2.884 and -11.090, respectively). While M176I nsSNPs had a neutral effect on CRISP2 protein with a threshold value greater than -2.5 (0.187), as shown in Table 1.

PolyPhen-2

Using PolyPhen to predict the possible effect of an amino acid substitution on the structure and hence the function of a protein, the L59V and C196R nsSNPs were predicted to have a potentially harmful impact on the CRISP2 protein, with scores of (0.988 and 0.999)

respectively. Table 1 shows that M176I was expected to be a benign nsSNP.

PHD‑SNP

The PhD-SNP, together with the reliability index score, is used to predict whether a novel phenotype resulting from an SNP is a neutral polymorphism or is related to genetic disorders in humans. As shown in Table 1, only C196R nsSNPs were associated with disease, whereas L59V and M176I nsSNPs were neutral.

Table 1. Consequences of SNPs on CRISP2 protein structure and function have been predicted by PROVEAN, PolyPhen2, and PhD-SNP

SNP ID	Mutant	PROVEAN score	Prediction	PolyPhen-2 Score	Prediction	PhD-SNP RI	Prediction	p- value
rs1765509750	L56V	-2.884	deleterious	0.988	probably damaging	4	Neutral	0.676
rs533319863	M176I	0.187	Neutral	0.000	Benign	2	Neutral	0.902
rs36069724	C196R	-11.090	deleterious	0.999	probably damaging	8	Disease	0.846

Prediction of SNPs effects on *CRISP2* protein stability by using I‑Mutant 2.0 server: A protein's biomolecular activities, functions, and regulation are generally greatly influenced by its structural stability. A decline in protein stability promotes protein dysfunction by causing protein breakdown, misfolding, and aggregation. Deleterious ns SNPs can change a protein's hydrophobicity and geometric constraints and disrupt hydrogen bonds and other bridges, which can affect the protein's ability to function normally (13,17). The native CRISP2 protein's sequence, or tertiary structure, was used to assess the influence of three mutations on the free energy change value (DDG) and reliability index (RI) and also to anticipate whether or not they will increase or decrease the stability of the CRISP2 protein as a consequence of a single point mutation.When compared to the natural protein, all of the mutant proteins (L56V, M176I, and C196R) revealed differing RI and DDG values, based on the results of the CRISP2 protein stability investigation. And all mutants' protein have decreased molecular stability due to the substitution of deleterious amino acids. DDG values were (-1.70 kcal/mol, -0.54 kcal/mol, and -0.22 kcal/mol, respectively). While M176I and C196R were neutral, L56V was found to be typically unstable. Moreover, as shown in Table 2, the RI values of these mutations were (8, 5, and 2), indicating that they reduced the stability of the CRISP2 protein. When compared to M176I and C196R, L56V has the highest reliability. As a result, these polymorphisms may harm the CRISP2 protein severely by lowering protein stability, which could lead to an increase in protein aggregation, misfolding, and degradation.

Table 2. Effect of SNPs on protein stability predicted by 1- Mutant v 2.0

SNP ID	Mutant	DDG Value Kcal /mol	RI values Score (0-10)	Stability Prediction	p- value
rs1765509750	L56V	-1.70	8	Decrease	0.676
rs533319863	M176I	-0.54	5	Decrease	0.573
rs36069724	C196R	-0.22	2	Decrease	0.498

According to studies, the difference in size between the native and mutant residues may cause the hydrogen bond to rupture, which might destabilize the local structure and packing. The original protein's ionic interactions will also be altered by the charge difference, leading to a loss of interactions with other molecules. For the purpose of spotting structural changes in a protein, the point mutation sites in the secondary structures are crucial. Alpha helices and beta strands are stabilized by hydrogen bonds (46, 36, 26).

While mutations in turns or loops have no effect, those in the alpha helix and beta sheet regions of the protein have a negative impact on the protein's structural integrity (27, 38). Leucine is a hydrophilic amino acid with a high hydrogen-binding capacity. When valine, a hydrophobic amino acid with distinctive physicochemical properties, was substituted for L56V, it could result in a reduction in hydrogen bonding in the L56V mutant. Proteins typically have polar amino acids in exposed locations, and any change in this region interferes with the protein's functionality (39). The C196R polymorphism includes the loss of a tightly conserved cysteine in the construction of one of the two crossed disulfide bonds in the hinge region bridging the CAP and CRISP domains in CRISPs (16, 15). The loss of this link would most likely lead to the hinge's hyper-flexibility, which could adversely affect how proteins interact with their binding partners. When a hydrophilic and positively charged arginine residue is swapped out for a non-polar cysteine residue, the cysteine-cysteine interaction of the native protein is hindered. The mutant residue has a larger and more hydrophilic molecular structure. The disulfide bond, which is necessary for the interaction of the CRISP2 protein, would be broken due to the difference in size and hydrophobicity between the original and mutant proteins. This mutation's detrimental nature is suggested by the expectation that it will render the CRISP2 protein ineffective. It's possible that the destabilization of the hinge region is what's causing the decline in protein interaction and rise in degradation susceptibility. It has been established through evolutionary stability studies and mutational resistance of protein-coding genes that arginine, leucine, and serine are the three main amino acids determining protein stability in mutants (28). Our investigations of stability shift in silico revealed that these three mutations (L56V, M176I, and C196R) had both structural and functional impacts on the mutant protein. These mutations may therefore be found in alpha helices and beta turns, where they might interfere with hydrogen bond formation and destabilize the protein. In the case of L59V or M176I, leucine and methionine may be involved in hydrogen bond formation. When compared to the natural structure, their replacement with valine and isoleucine may result in a decrease in the number of hydrogen bonds, lowering protein stability and increasing the likelihood of protein interaction with other molecules. Therefore, to confirm the correctness of our prediction techniques, the mutants might be subjected to a study of the protein's behaviour. In essence, the quantity of disulfide bonds formed in the native protein is dynamic and determines the protein's flexibility; these interactions are regarded as a crucial component in protein stability (19). Unlike the mutant C196R, which only contains 15, the original CRISP2 protein possesses 16 highly conserved disulfide bonds. This reduction can be a sign that protein structure and function have been negatively impacted. Protein stability would be reduced, and protein destabilization would result from the mutation C196R, which replaces the cysteine residue at position 196 with the hydrophilic amino acid arginine. The neutrally charged cysteine is smaller than the positively charged arginine. The disulfide bridge at the CRISPs' hinge region, which connects the CAP and CRISP domains, would be disrupted by the size and charge mismatch between the native and mutant residues.

Ionic interactions with the native protein would be impacted by the charge change as well. The differences in secondary structural components between the native and mutant proteins, as well as the stability difference brought on by the mutations, need to be further explored. To accomplish this, other tools need to be applied to examine these differences.

REFFERENCES

1. Adzhubei, I., Jordan, DM. and SR. Sunyaev. 2013. Predicting functional effect of human missense mutations using PolyPhen-2. Current Protocols in Human Genetics. 76(1):7–20

2. Adzhubei, I. A., Schmidt, S., Peshkin, L., Ramensky, V. E., Gerasimova, A., Bork, P., Kondrashov, A. S., & Sunyaev, S. R. 2010. A method and server for predicting damaging missense mutations. Nature methods, 7(4), 248–249.

3. Adzhubei, I.A, Sunyaev, S., Do, R., Balick, D., Li, H. and D. Reich. 2015. No evidence

that selection has been less effective at removing deleterious mutations in Europeans than in Africans. Nature Genetics. 47(2):126-31.

4. Akhtar, M., Jamal, T., Jamal, H., Din, J.U., Jamal, M., Arif, M., Arshad, M. and F. Jalil. 2019. Identification of most damaging nsSNPs in human CCR6 gene: In silico analyses. International Journal of Immunogenetics. 46: 459–471.

5. Akhta.r, M., Khan, S., Ali, Y., Haider, S., Din, J.U., Islam, Z.-U. and F. Jalil. 2020. Association study of CCR6 rs3093024 with Rheumatoid Arthritis in a Pakistani cohort. Saudi Journal of Biological Sciences. 27: 3354–3358.

6. Bava, K.A., Gromiha, M.M., Uedaira, H., Kitajima, K. and A. Sarai. 2004. ProTherm, version 4.0: thermodynamic database for proteins and mutants. Nucleic Acids Research. 32:

7. Bhagwat, M. 2010. Searching NCBI's dbSNP database. Current Protocols in Bioinformatics. Chapter 1: Unit 1: 19.

8. Capriotti, E., Calabrese, R. and R. Casadio. 2006. Predicting the insurgence of human genetic diseases associated to single point protein mutations with support vector machines and evolutionary information. Bioinformatics. 22(22):2729–34.

9. Capriotti, E., Fariselli, P. and R. Casadio. 2005. I-Mutant2. 0: predicting stability changes upon mutation from the protein sequence or structure. Nucleic Acids Research. 33(2): 306–10.

10. Choi, Y. and A.P. Chan. 2015. Provean web server: a tool to predict the functional effect of amino acid substitutions and indels. Bioinformatics. 31(16): 2745-2747

11. Choi, Y., Sims, GE., Murphy, S., Miller, JR. and AP. Chan. 2012. Predicting the functional effect of amino acid substitutions and indels. Open access journal. 7(10): e46688.

12. Dakal, TC., Kala, D., Dhiman, G., Yadav, V., Krokhotin, A. NV. and Dokholyan. 2017. Predicting the functional consequences of non-synonymous single nucleotide polymorphisms in IL8 gene. Scientific Reports. 7 (1):6525.

13. Deller, MC., Kong, L. and B. Rupp. 2016. Protein stability: a crystallographer's perspective. Acta Crystallographica Section. 72(2):72–95.

14. Gibbs, G.M., Lo, J.C.Y., Nixon, B., Jamsai, D., O'Connor, AE., Rijal, S., Sanchez-Partida, LG., Hearn, MTW., Bianco, DM. and MK. O'Bryan. 2010. Glioma pathogenesis-related 1-like 1 is testis enriched, dynamically modified, and redistributed during male germ cell maturation and has a potential role in sperm-oocyte binding. Endocrinology. 151:2331-2342

15. Gibbs, GM., Scanlon, MJ., Swarbrick, J., Curtis, S., Gallant, E., Dulhunty, AF. and MK. O'Bryan. 2006. The cysteine-rich secretory protein domain of Tpx-1 is related to ion channel toxins and regulates ryanodine receptor Ca2+ signaling. Journal of Biological Chemistry. 281:4156–4163

16. Guo, M., Teng, M., Niu, L., Liu, Q., Huang, Q. and Q. Hao. 2005. Crystal structure of the cysteine-rich secretory protein stecrisp reveals that the cysteine-rich domain has a K channel inhibitor-like fold. Journal of Biological Chemistry.280:12405-12412

17. Hossain, MS., Roy, AS. and Islam, MS. 2020. In Silico Analysis Predicting Effects of Deleterious SNPs of Human RASSF5 Gene on its Structure and Functions. Scientific Reports.10(1):14542.

18. Jamsai, D., Bianco, D. M., Smith, SJ., Merriner, DJ., Ly-Huynh, JD., Herlihy, A., Niranjan, B., Gibbs, GM. and O'Bryan MK. 2008. Characterization of gametogenetin 1 (GGN1) and its potential role in male fertility through the interaction with the ion channel regulator, cysteine-rich secretory protein 2 (CRISP2) in the sperm tail. Reproduction. 135:751–759

19. Jelesarov, I., and Karshikoff, A. 2009. Defining the role of salt bridges in protein stability. Methods in molecular biology, 490, 227-60.

20. Kellogg, E.H., Leaver-Fay, A. and D. Baker. 2011. Role of conformational sampling in computing mutation-induced changes in protein structure and stability. Proteins: Structure, Function, and Bioinformatics. 79(3):830–838.

21. Kliesch, S. 2014. Diagnosis of Male Infertility: Diagnostic Work-up of the Infertile Man. European Urology Supplements. 13(4): 73-82.

22. Knottnerus, S.J.G., Nijmeijer, S.C.M., IJlst, L., teBrinke, H., van Vlies, N. and FA. Wijburg. 2017. Prediction of phenotypic severity inmucopolysaccharidosis type IIIA. Annals of Neurology. 82:686–696
23. Koppers, A. J., Reddy, T. and MK. O'Bryan. 2011. The role of cysteine‐rich secretory proteins in male fertility. Asian Journal of Andrology. 13: 111–7
24. Krausz, C.G. and D.T. Carrell. 2014. Advances in understanding the genetics underlying male infertility and evolving diagnostic and treatment options. Andrology. 2:302-303
25. Lim, S., Kierzek, M., O'Connor, A. E., Brenker, C., Merriner, D. J., Okuda, H., Volpert, M., Gaikwad, A., Bianco, D., Potter, D., Prabhakar, R., Strünker, T., & O'Bryan, M. K. (2019). CRISP2 Is a Regulator of Multiple Aspects of Sperm Function and Male Fertility. Endocrinology, 160(4), 915–924.
26. Mohammed, I., Al-Awadi, S. & Saadedin, Sh. (2018). The Relationship between single nucleotide polymorphism of interleukin-10 gene promoter (-1082 A/G) with infection CHLAMYDIA TRACHOMATIS infertile Iraqi woman. Iraqi Journal of Agricultural Sciences. 49. 478-486. https://doi.org/10.36103/ijas.v49i3.119
27. Mosaeilhy, A., Mohamed, M.M., C, G.P., El Abd, H.S., Gamal, R., Zaki, O.K., & Zayed, H. (2017). Genotype-phenotype correlation in 18 Egyptian patients with glutaric acidemia type I. Metabolic Brain Disease, 32, 1417-1426.
28. Prosdocimi, F., Ortega, M. J. 2007. The codon usage of leucine, serine and arginine reveals evolutionary stability of proteomes and protein-coding genes. In: BrazSympos Bioinform. 149–159
29. Ramensky, V., Bork, P., and Sunyaev, S. 2002. Human non-synonymous SNPs: server and survey. Nucleic acids research, 30(17), 3894–3900.
30. Semagn, K., Babu, R., Hearne, S., and M. Olsen. 2014. Single nucleotide polymorphism genotyping using Kompetitive Allele Specific PCR (KASP): Overview of the technology and its application in crop improvement. Molecular Breeding. 33(1): 1-14
31. Sharma, A. 2017. Male Infertility; evidences, risk Factors, causes, diagnosis and management in human. Annals of Clinical and laboratory Research. 5 (3): 188
32. Shastry B. S. (2009). SNPs: impact on gene function and phenotype. Methods in molecular biology, 578, 3–22.
33. Shorter, J. R., Odet, F., Aylor, D. L., Pan, W., Kao, C. Y., Fu, C. P., and R. W. Feathers. 2017. Male infertility is responsible for nearly half of the extinction observed in the mouse collaborative cross. Genetics. 206(2):557-572.
34. Sidhu, N.S., Schreiber, K., Propper, K., Becker, S., Uson, I., Sheldrick, G.M., Gartner, J., Kratznera, R. and Steinfelda, R. 2014. Structure of sulfamidase provides insight into the molecular pathology of mucopolysaccharidosis IIIA. Acta Crystallographica Section D. 70 (5):1321–1335
35. Skakkebaek, NE., Rajpert-De Meyts, E., Buck Louis, GM., et al. 2016. Male reproductive disorders and fertility trends: influences of environment and genetic susceptibility. Physiological Reviews. 96:55–97
36. Sneha, P., Ebrahimi, EA., Ghazala, SA., Thirumal Kumar, D., Siva, R., George Priya Doss, C. and H. Zayed. 2018a. Structural analysis of missense mutations in Galactokinase (GALK1) leading to Galactosemia type-2. Journal of Cellular Biochemistry.119:1–14.
37. Sneha, P., Thirumal Kumar, D., George Priya Doss, C., Siva, R. and H. Zayed. 2017b. Determining the role of missense mutations in the POU domain of HNF1A that reduce the DNA-binding affinity: a computational approach. PLoS One. 12: e0174953
38. Sneha, P., Zenith, TU., Abu Habib, US., Evangeline, J., Thirumal Kumar, D., George Priya Doss, C., Siva, R. and H. Zayed. 2018b. Impact of missense mutations in survival motor neuron protein (SMN1) leading to spinal muscular atrophy (SMA): a computational approach. Metabolic Brain Disease. 33(6):1823–1834
39. Sudhakar, N., George Priya Doss, C., Thirumal Kumar, D., Chakraborty, C., Anand, K., & Suresh, M. (2016). Deciphering the impact of somatic mutations in exon 20 and exon 9 of PIK3CA gene in breast tumors among Indian women through molecular

dynamics approach. Journal of Biomolecular Structure and Dynamics, 34, 29 – 41

40. Thoma, ME., McLain, AC., Louis, JF., King, RB., Trumble, AC., Sundaram, R., and Buck Louis, GM. 2013. Prevalence of infertility in the United States as estimated by the current duration approach and a traditional constructed approach. Fertility and Sterility. 99: 1324–31

41. Trofimova, N. S., Olkhovich, N.V. and Gorovenko, NG. 2014. Specificities of Sanfilippo a syndrome laboratory diagnostic. Biopolymers and Cell. 30(5):388–393

42. Tüttelmann, F., Ruckert, C. and A Röpke. 2018. Disorders of spermatogenesis: perspectives for novel genetic diagnostics after 20 years of unchanged routine. Journal of Medical Genetics. 30(1):2–20

43. Venselaar, H., TeBeek, TA., Kuipers, RK., Hekkelman, ML. and G .Vriend. 2010. Protein structure analysis of mutations causing inheritable diseases. An e-Science approach with life scientist friendly interfaces. BMC Bioinformatics. 11(1):548

44. Virtanen, HE., Jorgensen, N. and J. Toppari. 2017. Semen quality in the 21st century. Nature Reviews Urology. 14:120–30

45. Winters, BR. And TJ. Walsh. 2014. The epidemiology of male infertility. Urologic Clinics of North America Journals. 41(1):195–204

46. Zaki, O. K., Krishnamoorthy, N., El Abd, H. S., Harche, S. A., Mattar, R. A., Al Disi, R. S., Nofal, M. Y., El Bekay, R., Ahmed, K. A., George Priya Doss, C., & Zayed, H. 2017. Two patients with Canavan disease and structural modeling of a novel mutation. Metabolic brain disease, 32(1),171–177.