Journal of Genetic Engineering and Biotechnology 15 (2017) 419–422

HOSTED BY

Contents lists available at ScienceDirect

Journal of Genetic Engineering and Biotechnology

journal homepage: www.elsevier.com/locate/jgeb

Original Article

Computational study of putative functional variants in human kisspeptin

Esmaeil Samadian [a,b,1], Roghaye Gharaei [a,1], Abasalt Hosseinzadeh Colagar [b,*], Hoda Sohrabi [c,*]

[a] Laboratory Sciences Research Center, Golestan University Of Medical Sciences, Gorgan, Iran
[b] Department of Molecular and Cell Biology, Faculty of Basic Sciences, University of Mazandaran, Babolsar, Iran
[c] Department of Medical Biotechnology, School of Advanced Technologies in Medicine, Golestan University Of Medical Sciences, Gorgan, Iran

ARTICLE INFO

Article history:
Received 11 February 2017
Received in revised form 23 April 2017
Accepted 24 July 2017
Available online 2 August 2017

Keywords:
Kisspeptin
Non-synonymous variation
Functional effect

ABSTRACT

Non-synonymous single nucleotide polymorphisms (nsSNPs) are a type of genetic mutations that result in amino acid substitution of the encoded proteins that may potentially affect its function and phenotype. An *In Silico* assay has been carried out by using bioinformatics prediction tools to identify nsSNPs which are responsible for important disorders in human kisspeptin (KISS1) gene. In this study, for the first time, KISS1 amino acid changes were discovered by tBlastn for EST database. A list of nsSNPs in human KISS1 gene from dbSNP, dbEST and UniProt databases were prepared. Computational analysis was performed using SIFT (Sorting Intolerant From Tolerant) and PolyPhen (Polymorphism Phenotyping) programs. Of the total 92 nsSNPs, 20 were found to be damaged by both servers. Six nsSNPs (P97L, G122R, W114C, R92C, R120H and N115K) are predicted with the highest damaging scores (SIFT = 0, PolyPhen = 1). These intolerant changes may suggest their functional significance in critical regions which may affect the function and stability of KISS1 protein. Identifying these nsSNPs among the thousands of them make an opportunity to screen only those predicted deleterious by programs.

1. Introduction

KISS1 is a tumor suppressor and antimetastatic gene and its Protein precursor, Kisspeptin, is implicated in cancer invasion, migration, metastasis and angiogenesis [1,2]. This is proteolysed to fragments of various lengths, including Kisspeptin-54, kisspeptin-10, kisspeptin-13 and kisspeptin-14 [3]. This peptide hormone binds to G-protein-coupled kisspeptin receptor (KISS1R). Kisspeptin-KISS1R signaling has a critical role in various physiological and pathophysiological processes such as urogenital system and reproductive function. Puberty is regulated by the maturation of kisspeptin neurons and by interactions between kisspeptins and leptin [4,5]. Studies showed that KISS1 had high expression in the cell cytoplasm of early-stage colorectal cancer cells While KISS1 expression was significantly higher in tumors with extrathyroidal invasion and advanced stage [6,7]. Unlike the recent study, expression level was considerably lower in metastasis tissues of lymph node (LN), brain and gallbladder adenocarcinoma compared with their primary tumor tissues. So, these studies demonstrates significant correlations between KISS1 expression and metastasis [8,9]. In addition, protein expression of both KISS1/KISS1R in the cancerous tissues compared with noncancerous tissue adjacent to the breast tumor has been increased [10]. Knocking down KISS1 resulted in increased invasion and migration of colorectal cancer cells [11]. Also, kisspeptin suppressed metastasis of urogenital carcinoma and reduced metastasis in malignant melanomas and breast cancer cell lines by inhibition of cellular chemotaxis and invasion [5,12].

Human KISS1 (NP_002247) mRNA encodes a protein of 138 amino acids with a predicted molecular mass of approximately 14.7 kDa. It is located at chromosome 1q32 [13].

Most human genetic variation is represented by single-nucleotide polymorphisms (SNPs) [14]. Non-synonymous single nucleotide polymorphism (nsSNP) is a type of genetic mutation that causes a single amino acid substitution (AAS) in a protein sequence [15]. nsSNPs affect gene regulation by altering DNA and transcriptional binding factors and the maintenance of the structural integrity of cells and tissues. Also, nsSNPs affect the functional roles of proteins involved in signal transduction of visual, hormonal, and other stimulants, subsequently altering the carrier's phenotype [16,17].

Peer review under responsibility of National Research Center, Egypt.
* Corresponding authors.
E-mail addresses: acolagar@yahoo.com (A. Hosseinzadeh Colagar), hsohrabi65@yahoo.com (H. Sohrabi).
[1] These authors contributed equally to this work.

http://dx.doi.org/10.1016/j.jgeb.2017.07.007
1687-157X/

Predicting the deleterious nsSNPs for a gene by *In Silico* tools is very useful, because screening all of the functionally important nsSNPs by experimental analysis are relatively hard. In this study, nsSNPs of KISS1 gene that potentially affect its protein function were predicted.

2. Materials and methods

2.1. Data mining

The nsSNPs and their related protein sequences of KISS1 gene were extracted from the National Center for Biotechnology Information (NCBI) database of SNPs, dbSNP (http://www.ncbi.nlm.nih.gov/projects/SNP) and UniProt (http://www.UniProt.org) for our computational analysis. To further screen for amino acid substitutions in KISS1, the amino acid sequence of KISS1_(NP_002247) against the most recent EST database (build 130) was compared using the tBlastn algorithm (http://www.ncbi.nlm.nih.gov/BLAST/) [18–20]. In total, 93 KISS1 or KISS1-related sequences were assessed. Only those variants that were observed more than once were scored, as possible amino acid substitutions and the rest were not included in this study; nonsense (∗) or ambiguous (B, Z, X, etc.) amino acid substitutions were also not considered for further analysis.

2.2. Analysis of deleterious nsSNPs by SIFT

The Sorting Intolerant from Tolerant (SIFT) algorithm predicts the effect of coding variants on protein function. SIFT is one of the standard tools for characterizing missense variation and predicts the effects of all possible substitutions at each position in the protein sequence by using sequence homology [15,21]. To detect the deleterious coding nsSNPs, we used SIFT and submitted the query in the form of either SNP identifier (ID)s or as protein sequences. SIFT analysis was carried out by allowing the algorithm to search for homologous sequences and using the default settings (SWISS-PROT 45 and TrEMBL 28 databases, median conservation score 3.00, remove sequences >90% identical to query sequence). The principle of this program is that it uses multiple sequence alignment conservation approach and calculates normalized probabilities for every possible substitution. Then, assigns scores to each residue, ranging from 0 to 1. Scores close to 0 indicate evolutionary conservation and is damaging to the protein, while scores close to 1 indicate that the substitution is tolerated by protein. The amino acid change is predicted damaging if the score is $\leq$0.05, and tolerated if the score is >0.05. SIFT scores <0.05 are predicted to be intolerant or deleterious amino acid substitutions, whereas scores >0.05 are considered tolerant.

2.3. Impact of an amino acid substitution predicted by PolyPhen

Polymorphism Phenotyping is an automatic tool for prediction of the possible impact of an amino acid substitution on the structure and function of a human protein [14]. The prediction is based on a number of sequence, phylogenetic, and structural features characterizing the substitution [13]. PolyPhen server accepts input options such as protein sequence, SWALL database ID or accession number along with sequence position with two amino acid alterations. The input query was submitted in the form of protein sequence together with the positions of the substitution (native) and substituting amino acids (mutant). The PolyPhen estimates sensitivity, specificity and calculate the PSIC (position-specific independent count) scores of the both variants. The PolyPhen also computes the PSIC scores difference between two variants. PolyPhen scores were determined as "probably damaging" (0.96–1), "possibly damaging" (0.71–0.95), "benign" (0.31–0.7) and "unknown" (0.00–0.3). If no prediction can be made due to a lack of data then the outcome is reported as "unknown".

3. Results

The human KISS1 gene retrieved from dbSNP, dbEST and UniProt databases contained a total of 92 missense SNPs and their corresponding information including allele changes, amino acid substitutions and substitution positions are listed in Table S1 and S2 (supplementary data).

The protein sequences of all 92 nsSNPs were submitted separately to the SIFT program to check its tolerance index. Out of 92 nsSNPs, 43 (46.73%) were found to be deleterious with the tolerance index score of <0.05 as shown in Table 1. These variants could affect the protein function in the KISS1 gene and the remaining 49 (53.26%) nsSNPs were analyzed to be tolerant in KISS1.

In PolyPhen analysis, protein sequence with mutational position and amino acid variants related to all the retrieved nsSNPs were submitted as inputs to PolyPhen server. Among the 92 variants submitted to the PolyPhen, 32 (34.78%) of these variants were predicted to be probably damaging for KISS1 protein, 26 (28.26%) to be possibly damaging, and 34 (36.95%) to be benign substitutions.

It is to be noted that among 43 nsSNPs that were observed to be deleterious by SIFT program 20 (46.51%) were also predicted to be highly damaging by PolyPhen, 13 (30.23%) possibly damaging and 10 (23.25%) benign.

Moreover, 4 variants (E20K, Q36R, P81R and N115K) were obtained from UniProt and analyzed using the servers. Three of them were exited in dbSNP and a remaining variant was found in dbEST database. The results of their analysis is displayed in Table 2.

4. Discussion

KISS1 gene in human encodes a group of fragment peptides which are called kisspeptins. They are C-terminally amidated peptide products, including KP- 10, KP-13, KP-14 and KP-54 and expressed in the hypothalamus, gonads, placenta, liver and pancreas [22]. Kisspeptins are essential for reproductive function and regulate puberty. Additionally, recent investigations have suggested some roles of Kisspeptin signalling in the suppression of metastasis with a variety of cancers such as colon, breast, brain, lymph node, thyroid, gall bladder carcinoma [1,7,9,23,24].

Recent studies indicated several functions of KISS1 in biological processes, though, the functional domains of the protein are not completely elucidated. Also, there is no 3 Dimensional structure for KISS1, up to now. But, the data were taken from several protein data sources showed the presence of a signal peptide between 1–19 amino acids (a.a.). Signal sequences are N-terminal extensions of newly synthesized secretory and membrane proteins. They are on average 16 to 30 amino acid residues in length comprising a characteristic tripartite structure: (1) a hydrophilic, usually positively charged n-region, (2) a central hydrophobic h-region and (3) a c-region with the cleavage site for signal peptidase [25]. Our analysis by polyphen software revealed a point mutation in the signal peptide-encoding KISS1 gene which can be probably damaging. Mutation of signal peptide at amino acid 18 (Phe – Cys) may affect cleavage of signal peptide and subsequently protein maturation.

Metastin (Kisspeptin 54) functions as the endogenous ligand of the G-protein coupled receptor GPR54. Activation of the receptor inhibits cell proliferation and cell migration, key characteristics of tumor metastasis [26]. Our *In Silico* analysis of the human KISS1

Table 1
Deleterious amino acid substitutions predicted by both SIFT and PolyPhen.

#rs ID/Amino acid change	SIFT score	PolyPhen score	#rs ID/Amino acid change	SIFT score	PolyPhen score
rs187437980/R [Arg] 123 W [Trp]	0.00	0.999	rs776352326/W [Trp] 114C [Cys]	0.00	1.000
rs532969667/G [Gly] 35 S [Ser]	0.03	0.999	rs780032947/R [Arg] 92C [Cys]	0.00	1.000
rs545669513/P [Pro] 97 L [Leu]	0.00	1.000	R [Arg] 120 H [His]	0.00	1.000
rs746488000/G [Gly] 79 W [Trp]	0.01	0.991	T [Thr] 34 Q [Gln]	0.00	0.998
rs755431201/A [Ala] 96 V [Val]	0.01	0.980	N [Asp] 115 K [Lys]	0.00	1.000
rs756429135/R [Arg] 137 Q [Gln]	0.00	0.999	S [Ser] 116 P[Pro]	0.02	0.999
rs760447526/R [Arg] 120 S [Ser]	0.00	0.988	E [Glu] 20 S [Ser]	0.00	0.984
rs764632722/R [Arg] 67 Q [Gln]	0.00	0.989	F [Phe] 18 C [Cys]	0.00	0.998
rs771480486/P [Pro] 95 R [Arg]	0.01	0.998	P [Pro] 21 H [His]	0.03	0.994
rs772730074/G [Gly] 122 R [Arg]	0.00	1.000	V [Val] 28 C [Cys]	0.01	0.997

Table 2
Analysis of natural variants collected from UniProt database.

#rs ID/ amino acid Substitution	SIFT prediction	SIFT score	PolyPhen prediction	PolyPhen score
N [Asp] 115 K [Lys]	Damaging	0.00	Probably damaging	1.000
rs35431622/Q [Gln] 36 R[Arg]	Tolerated	0.62	Benign	0.100
rs4889/P [Pro] 81 R [Arg]	Tolerated	1.00	Possibly damaging	0.524
rs12998/E [Glu] 20 K [Lys]	Damaging	0.00	Possibly damaging	0.732

mutations showed the functional importance of changes as predicted to be structurally deleterious. These mutations may change total electricity charge of protein and consequently alter its function. Since Metastin functions as a cell proliferation and cell migration inhibitor, vulnerability and resistance of people to different cancers may be influenced by these mutations.

Also, according to UniProt database, essential amino acids for receptor binding and receptor activation are located between 112 and 120. Five point mutations (W114C, N115K, S116P, R120H and R120S) which are predicted to be deleterious may affect protein binding to the GPR54 and subsequently critical roles of Kisspeptin-KISS1R signaling in various physiological and pathophysiological processes.

Discovery of human KISS1 amino acid changes by tBlastn for EST database done for the first time in our study, indicated that novel variations may potentially affect the corresponding protein function, as mentioned in the Table 1. Interestingly, substitution N115K (SIFT score = 0, PolyPhen score = 1) associated with HH13, was also predicted by our analysis highly damaging, as can be seen from Table 2. It suggests that four neighbor mutations of this region, may affect the susceptibility to HH13.

There is no direct approach of evaluating the accuracy of these predictions made by SIFT and PolyPhen, as it is possible that the algorithms used different data sets. So, even if the predictions by these programs were not completely consistent for this subset of mutants, these KISS1 variants still should be regarded as candidates for SNP screening. Furthermore, molecular modeling and experimental investigations can be used to confirm our results [27].

5. Conclusions

It could be concluded that Identifying intolerant substitutions among the lots of them make it possible to screen only those predicted deleterious by programs. It seems that these variants may be situated within a functionally important region and may affect the stability, folding of KISS1 protein and eventually people's cancer susceptibility.

Acknowledgments

Our project has got an approval from the Golestan University of Medical Sciences, Golestan, Iran. We acknowledge all cancer researchers worldwide and also, our friend, Semira Kheiri for her helpful contribution.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.jgeb.2017.07.007.

References

[1] Kostakis ID, Agrogiannis G, Vaiopoulos AG, Mylona E, Patsouris E, Kouraklis G, et al. A clinicopathological analysis of KISS1 and KISS1R expression in colorectal cancer. 2015;123(7):629–37.
[2] Kostakis ID, Agrogiannis G, Vaiopoulos AG, Mylona E, Patsouris E, Kouraklis G, et al. KISS1 expression in colorectal cancer. 2013;121(10):1004–10.
[3] Kotani M, Detheux M, Vandenbogaerde A, Communi D, Vanderwinden J-M, Le Poul E, et al. The Metastasis Suppressor Gene KiSS-1 Encodes Kisspeptins, the Natural Ligands of the Orphan G Protein-coupled Receptor GPR54. J Biol Chem 2001;276(37):34631–6.
[4] Tng EL. Kisspeptin signalling and its roles in humans. Singapore Med J 2015;56 (12):649–56.
[5] Wahab F, Atika B, Shahab M, Behr R. Kisspeptin signalling in the physiology and pathophysiology of the urogenital system. Nat Rev Urol 2016;13(1):21–32.
[6] Gupta PB, Onder TT, Jiang G, Tao K, Kuperwasser C, Weinberg R a, et al. Identification of selective inhibitors of cancer stem cells by high-throughput screening. Cell 2009;138(4):645–59.
[7] Savvidis C, Papaoiconomou E, Petraki C, Msaouel P, Koutsilieris M. The role of KISS1/KISS1R system in tumor growth and invasion of differentiated thyroid cancer. Anticancer Res 2015;35(2):819–26.
[8] Shin W-J, Cho Y-A, Kang K-R, Kim J-H, Hong S-D, Lee J-I, et al. KiSS-1 expression in oral squamous cell carcinoma and its prognostic significance 2016;124 (4):291–8.
[9] Wang W, Yang Z, Liu J, Yang L, Yang X, Fu X. Overexpression of MTA1 and loss of KAI-1 and KiSS-1 expressions are associated with invasion, metastasis, and poor-prognosis of gallbladder adenocarcinoma. Tumori 2014;100(6):667–74.
[10] Jarzabek K, Koda M, Kozlowski L, Milewski R, Wolczynski S. Immunohistochemical study of KiSS1 and KiSS1R expression in human primary breast cancer: association with breast cancer receptor status, proliferation markers and clinicopathological features. Histol Histopathol 2015;30(6):715–23.
[11] Ji K, Ye L, Ruge F, Hargest R, Mason MD, Jiang WG. Implication of metastasis suppressor gene, Kiss-1 and its receptor Kiss-1R in colorectal cancer. BMC Cancer 2014;14:723.
[12] Quevedo EGC, Aguilar GMM, Aguilar LAJ, Rubio SAG, Martínez SEF, Rodríguez IPD, et al. Polymorphisms rs12998 and rs5780218 in KiSS1 suppressor metastasis gene in Mexican patients with breast cancer. Dis Markers 2015;2015:365845.
[13] West A, Vojta PJ, Welch DR, Weissman BE. Chromosome localization and genomic structure of the KiSS-1 metastasis suppressor gene (KISS1). Genomics 1998;54(1):145–8.
[14] Adzhubei I, Jordan DM, Sunyaev SR. Predicting functional affect of human missense mutations using PolyPhen-2. Curr Protoc Hum Genet 2013;7(7):20.

[15] Kumar P, Henikoff S, Ng PC. Predicting the affects of coding non-synonymous variants on protein function using the SIFT algorithm. Nat Protoc 2009;4 (7):1073–81.
[16] Rajasekaran R, Sudandiradoss C, Doss CGP, Sethumadhavan R. Identification and in silico analysis of functional SNPs of the BRCA1 gene. Genomics 2007;90 (4):447–52.
[17] Dryja TP, McGee TL, Hahn LB, Cowley GS, Olsson JE, Reichel E, et al. Mutations within the rhodopsin gene in patients with autosomal dominant retinitis pigmentosa. N Engl J Med 1990;323(19):1302–7.
[18] Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. J Mol Biol 1990;215(3):403–10.
[19] Gish W, States DJ. Identification of protein coding regions by database similarity search. Nat Genet 1993;3(3):266–72.
[20] Boguski MS, Lowe TMJ, Tolstoshev CM. DbEST — database for "expressed sequence tags". Nat Genet 1993;4(4):332–3.
[21] Sim N-L, Kumar P, Hu J, Henikoff S, Schneider G, Ng PC. SIFT web server: predicting affects of amino acid substitutions on proteins. Nucl Acids Res 2012;40(Web Server issue):W452–7.
[22] Prabhu VV, Sakthivel KM, Guruvayoorappan C. Kisspeptins (KiSS-1): essential players in suppressing tumor metastasis. Asian Pac J Cancer Prev 2013;14 (11):6215–20.
[23] Ulasov IV, Kaverina NV, Pytel P, Thaci B, Liu F, Hurst DR, et al. Clinical significance of KISS1 protein expression for brain invasion and metastasis. Cancer 2012;118(8):2096–105.
[24] Beug H. Breast cancer stem cells: eradication by differentiation therapy? Cell 2009;138(4):623–5.
[25] Kapp K, Schrempf S, Lemberg MK, et al. Post-targeting functions of signal peptides. In: Madame Curie Bioscience Database. Austin (TX): Landes Bioscience; 2000–2013. Available from: <https://www.ncbi.nlm.nih.gov/books/NBK6322/>.
[26] Ohtaki T, Shintani Y, Honda S, Matsumoto H, Hori A, Kanehashi K, et al. Metastasis suppressor gene KiSS-1 encodes peptide ligand of a G-protein-coupled receptor. Nature 2001;411(6837):613–7.
[27] Naderi M, Gharaei R, Soleymani-Nejadian E, Samadian E. In Silico survey of functional coding variants in human AEG-1 gene. Egypt J Med Hum Genet 2013;14(4):419–22.