Journal of Animal and Veterinary Advances 10 (7): 928-931, 2011
ISSN: 1680-5593

A Comprehensive *in silico* Analysis of Functional and Structural Impact SNPS in the *MC1R* Gene

[1]Chuan-Sheng Zhang, [2]Li-Ying Geng, [1]Zheng-Zhu Liu, [1]Zhi-Xin Fu, [1]Yuan-Fang Gong, [1]Min-Shan Feng, [1]Chen Juan, [1]Qing-Hui Jia, [1]Qiu-Yue Wang, [1]Xie-Rong Liu and [1]Su-Min Pan
[1]Department of Animal Science, [2]Department of Life Science,
Hebei Normal University of Science and Technology, 066600 Changli, China

Abstract: The aim of this study was to analyze the genetic variation that can alter the expression and the function of the *MC1R* gene using computational methods. Out of the total 222 SNPs, 22 were found to be non-synonymous (ns) SNPs, 9 were found in the 5' upstream and 7 were found in the 3'UTR. It was found that these 6 nsSNPs rs3212366, rs11547464, rs1805009, rs1805008 and rs1805007 were damaging by both the SIFT and the Poly Phen servers. We identified, a mutation from Phe to Leu at positions 196 (rs3212366) on the surface of the protein caused the greatest impact on stability. The rs3212362, rs3212379 and rs35025176 in the 5' upstream were suggested might change the *MC1R* gene expression levels by FastSNP. Based on the *in silico* search, the rs3212369 located in the putative Hsa-miR-421 target sequences might have some effect on *MC1R* gene expression.

Key words: MC1R, SNPs, SIFT, PolyPhen, FastSNP, China

INTRODUCTION

Single-Nucleotide Polymorphisms (SNPs) are the most frequent variation in the human genome, occurring once every several hundred base pairs throughout the genome. One of the main goals of SNP research is to understand the genetics of the human phenotype variation and especially the genetic basis of complex diseases (Collins *et al*., 1998). Non-synonymous coding SNPs (nsSNPs) comprise a group of SNPs that together with SNPs in regulatory regions are believed to affect gene regulation by altering DNA and transcriptional binding factors and then have the highest impact on phenotype in the human population (Bernig and Chanock, 2006). Recently, many evidences indicated that SNPs located at microRNA-binding sites of 3'UTR are likely to affect the expression of the microRNA target and may contribute to the susceptibility of humans to common diseases (Yu *et al*., 2007).

Melanocortin 1 receptor (MC1R), a Gs protein-coupled receptor expressed in melanocytes is a major determinant of skin pigmentation, phototype and cancer risk (Sturm, 2002). *MC1R* gene mutations greatly increase the risk of skin cancer has been demonstrated (Matichard *et al*., 2004). At present most mutations in the *MC1R* gene have been identified but only a few naturally occurring alleles have been functionally characterized. In this study, we perform a computational analysis of the SNPs in the *MC1R* gene to identify the deleterious nsSNPs that are likely to affect the function. In addition, the consequences would be the foundation for the preparation of the structure analysis for the native and mutant protein.

MATERIALS AND METHODS

The SNPs and their related protein sequences for the *MC1R* gene were retrieved from the dbSNP (Sherry *et al*., 2001). The deleterious coding nsSNPs were detected by the Web program SIFT (Ng and Henikoff, 2003) and PolyPhen (Ramensky *et al*., 2002). Substitutions at each position with normalized probabilities less than a chosen cutoff are predicted to be deleterious and those greater than or equal to the cutoff are predicted to be tolerated. The cutoff value in the SIFT program is a tolerance index of ≥ 0.05. The higher the tolerance index, the less functional impact a particular amino acid substitution is likely to have. The PSIC scores for each of the 2 variants were calculated. The higher the PSIC score difference is the functional impact a particular amino acid substitution is likely to have. Structural analysis was performed in order to evaluate and compare the stability of native and mutant structures. The 3-D model of the human *MC1R* gene was predicted using the CPH models 2.0

Corresponding Author: Li-Ying Geng, Department of Life Science, Hebei Normal University of Science and Technology, 066600 Changli, China

Server (Nielsen *et al*., 2010). The comparison between the resulting native and modeled structures was made by the calculation of the potential energy and RMSD values using the Swiss-Pdb Viewer (Guex and Peitsch, 1997). The web server FastSNP was used to predict the functional significance of nsSNPs, flanking regions SNPs and UTR SNPs of the *MC1R* gene (Yuan *et al*., 2006). The web service TFSearch was used to predict whether a noncoding SNP alters the transcription factor-binding site of a gene.

The score will be given by this server on the basis of levels of risk with a ranking of 0, 1, 2, 3, 4 or 5. This signifies the levels of no, very low, low, medium, high and very high effect, respectively. The miRanda software were used to predict 3'UTR microRNA target and analyze effects of SNPs on microRNA target (Griffiths-Jones *et al*., 2008).

RESULTS AND DISCUSSION

A total of 222 SNPs within human *MC1R* gene were retrieved from the dbSNP database (Sherry *et al*., 2001) of which 22 were nsSNPs, 10 were synonymous SNPs in coding regions and 16 were in flanking regions which comprise 9 SNPs in the 5' upstream and 7 SNPs in the 3'UTR. The conservation level of a particular position in a proteinwas determined by using a sequence homology-based tool, SIFT (Ng and Henikoff, 2003). Among the 22 nsSNPs, 12 were found to be deleterious according to SIFT having a tolerance index score of <0.05. Out of 12 deleterious nsSNPs, 8 showed a highly deleterious tolerance index score of 0.00 and 3 showed a tolerance index score of 0.03 followed by 1 nsSNPs with a tolerance index of 0.01. The results are shown in Table 1.

The structural levels of alteration were determined by applying the PolyPhen program (Ramensky *et al*., 2002). About 22 protein sequences of nsSNPs investigated were submitted as input to the PolyPhen server and the results are also shown in Table 1. It can be seen that of 22 nsSNPs, 11 were considered to be damaging. All 11 nsSNPs exhibited a PSIC score difference in the range 1.294-2.751. Hence, we could infer that the results obtained on the basis of sequence details (SIFT) were in good correlation with the results obtained for structural details (PolyPhen).

Table 1 shows these 5 nsSNPs rs3212366, rs11547464, rs1805009, rs1805008 and rs1805007 had a SIFT tolerance index of 0.00 and PSIC score difference ≥2.00. Hence, the mutations occurring with these 6 nsSNPs would be of prime importance in the identification of cancer caused by the *MC1R* gene according to SIFT and PolyPhen results.

Table 1: List of nsSNPs that were analysed by SIFT and PolyPhen

SNP ID	Nucleotide change	Amino acid change	Tolerance index	PSIC CD
rs3212366	C/T	F196L	0.00	2.724
rs11547464	A/G	R142H	0.00	2.552
rs1805009	C/G	D294H	0.00	2.505
rs1805008	C/T	R160W	0.00	2.751
rs1805007	C/G/T	R151C	0.00	2.737
rs1805006	A/C	D84E	0.00	1.986
rs1805005	G/T	V60L	0.00	-
rs1110400	C/T	I155T	0.01	1.390
rs3212365	C/G	V156L	0.03	1.298
rs2229617	A/G	G104L	0.03	1.759
rs12102534	C/T	T272M	0.03	1.294

Table 2: RMSD and total energy of native structure and mutant modeled structures

dbSNP ID	Amino acid change	RMSD between native and mutant structures	Total energy after minimization
rs1805007	R151C,arg151Cys	2.46 A°	-14533.371 KJ moL^{-1}
rs1805008	R160W,arg160Trp	1.99 A°	-14742.885 KJ moL^{-1}
rs3212366	F196L,phe196Leu	2.81 A°	-14486.227 KJ moL^{-1}
rs11547464	R142H,arg142His	2.37 A°	-14660.183 KJ moL^{-1}
rs1805009	D294H,asp294His	2.14 A°	-14718.856 KJ moL^{-1}

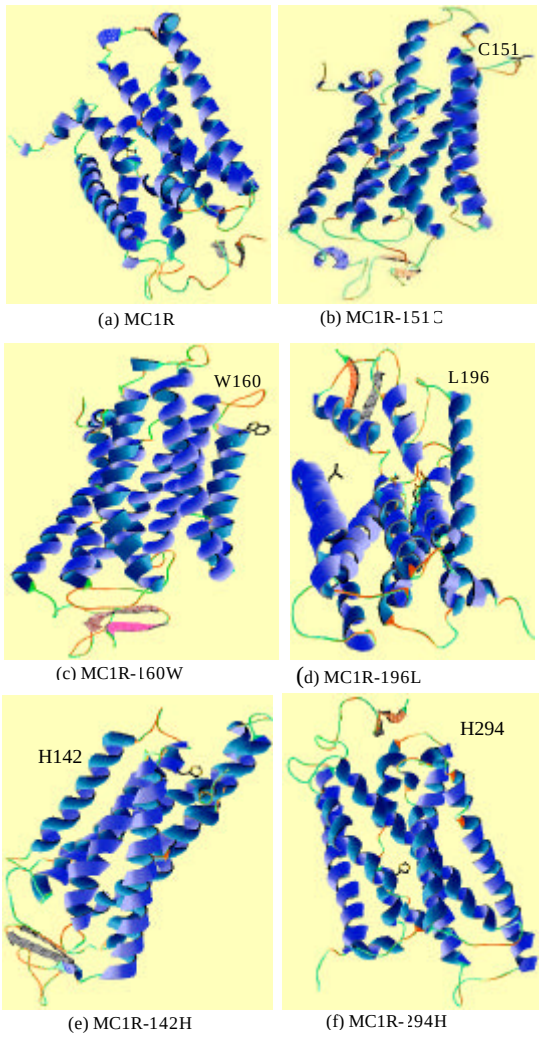
Total energy of native structure after energy minimization: -14882.880 KJ moL^{-1}

Out of common 6 nsSNPs predicted to be deleterious by SIFT and PolyPhen, 5 (rs3212366, rs11547464, rs1805009, rs1805008 and rs1805007) were mapped to the PDB MC1R native structure (NP_002377) and five mutant modeled structures and energy minimizations were performed by SPDBV (4.0) to get (MC1R, MC1R-152C, MC1R-160W, MC1R-196L, MC1R-142H and MC1R-294H, respectively).

The results are shown in Table 2 and Fig. 1a-f. These 5 mutant modeled structures showed an increase in energy in comparison with the native structure. Further, we identified a mutation from Phe to Leu at position 196 (rs3212366) on the surface of the protein caused the greatest impact on stability. The RMSD values between the native structure and the mutant modeled structures are ranging from 1.99- 2.81A°C. Because these values are high, we can suggest that these mutations may cause a significant change in the mutant structures with respect to the native protein structure.

According to the server FastSNP of 16 flanking regions SNPs of the *MC1R* gene, 3 in the 5' upstream, namely rs3212362, rs3212379 and rs35025176 were predicted to be damaging with a risk ranking of 1-3, totally. The results are shown in Table 3. However, this server did not predict any functional significance for the 3'UTR and hence, the results is further required to research.

Recently, a novel class of functional polymorphisms termed miRSNPs/polymorphisms in the 3'UTR was reported defined as a polymorphism present at a microRNA binding sites of functional genes that can affect gene expression by interfering with a micro RNA



(a) MC1R (b) MC1R-151C (c) MC1R-160W (d) MC1R-196L (e) MC1R-142H (f) MC1R-294H

Fig. 1: Superimposed structure of native *MC1R* gene with (a) mutant structure (b-f)

Table 3: List of SNPs predicted to be functionally significant by FastSNP

SNP ID	Nucleotide change	UTR position	Level of risk	Functional element change
rs3212362	A/G	5'UTR	Very low-medium (1-3)	Promoter/ regulatory region
rs3212379	C/T	5'UTR	Very low-medium (1-3)	Promoter/ regulatory region
rs35025176	-/G	5'UTR	Very low-medium (1-3)	Promoter/ regulatory region

Table 4: Prediction results of 3'UTR microRNA target the SNP

Target	Nucleotide change	Micro RNA	Mutation position
rs3212369	A/G	Hsa-miR-421	5'AAUGAUCUCUG AAAGUGUUGAAG || ||||| 3'CGCGGGUUAAU UACAG—ACAACUA || ||||| 5'AAUGAUCUCUGA AGGUG UUGAAG

function. A total of 41 microRNA targets were predicted by applying the miRanda software. A (A>G with ID rs3212369) within putative Hsa-miR-421 targets was found as shown in Table 4.

CONCLUSION

In this study, we investigated the functional and structural impact of SNPs in the skin cancer *MC1R* gene using computational prediction tools. Out of a total of 222 SNPs in the *MC1R* gene, 22 SNPs were found to be non-synonymous, 9 and 7 SNPs occurred in 5' upstream and 3'UTR, respectively. Out of 22 nsSNPs, 12 were found to be deleterious by SIFT and 11 were found to be damaging by the PolyPhen tool. A total of ten nsSNPs were found to be common in both the SIFT and the PolyPhen tool. The structural analysis results showed that the amino acid residue substitutions which had the greatest impact on the stability of the *MC1R* protein were mutations F196L (rs3212366). Among the three SNPs in the 5' upstream studied, 3 SNPs (rs3212362, rs3212379 and rs35025176) were found to be of functional significance. A mutation (A>G with ID rs3212369) in 3'UTR was predicted to affect Hsa-miR-421 by TargeScan. Based on the results, we conclude that these SNPs should be considered important candidates in causing diseases related to *MC1R* gene malfunction.

ACKNOWLEDGEMENTS

This research was supported by key projects in the National Science and Technology Pillar Program during the 11th 5 years plan period (No. 2008BADB2B04-8-4) and Natural Science Foundation of Hebei Province (no. C2008001308).

REFERENCES

Bernig, T. and S.J. Chanock, 2006. Challenges of SNP genotyping and genetic variation: Its future role in diagnosis and treatment of cancer. Expert Rev. Mol. Diagn., 6: 319-331.

Collins, F.S., L.D. Brooks and A. Chakravarti, 1998. A DNA polymorphism discovery resource for research on human genetic variations. Genome Res., 8: 1229-1231.

Griffiths-Jones, S., H.K. Saini, S. van Dongen and A.J. Enright, 2008. miRBase: Tools for microRNA genomics. Nucl. Acids Res., 36: D154-D158.

Guex, N. and M.C. Peitsch, 1997. SWISS-MODEL and the Swiss-PDBViewer: An environment for comparative protein modeling. Electrophoresis, 18: 2714-2723.

Matichard, E., P. Verpillat, R. Meziani, B. Gerard and V. Descamps *et al*., 2004. Melanocortin 1 receptor (MC1R) gene variants may increase the risk of melanoma in France independently of clinical risk factors and UV exposure. J. Med. Genet., 41: e13-e13.

Ng, P.C. and S. Henikoff, 2003. SIFT: Predicting amino acid changes that affect protein function. Nucl. Acids Res., 31: 3812-3814.

Nielsen, M., C. Lundegaard, O. Lund and T.N. Petersen, 2010. CPHmodels-3.0-remote homology modeling using structure-guided sequence profiles. Nucl. Acids Res., 38: W576-W581.

Ramensky, V., P. Bork and S. Sunyaev, 2002. Human non-synonymous SNPs: Server and survey. Nucl. Acids Res., 30: 3894-3900.

Sherry, S., M. Ward, M. Kholodov, J. Baker, L. Phan and E.M. Smigielski, 2001. dbSNP: The NCBI database of genetic variation. Nucl. Acids Res., 29: 308-311.

Sturm, R.A., 2002. Skin colour and skin cancer-MC1R, the genetic link. Melanoma Res., 12: 405-416.

Yu, Z., Z. Li, N. Jolicoeur, L. Zhang and Y. Fortin *et al*., 2007. Aberrant allele frequencies of the SNPs located in microRNA target sites are potentially associated with human cancers. Nucl. Acids Res., 35: 4535-4541.

Yuan, H.Y., J.J. Chiou, W.H. Tseng, C.H. Liu and C.K. Liu *et al*., 2006. FASTSNP: An always up-to-date and extendable service for SNP function analysis and prioritization. Nucl. Acids Res., 34: W635-W641.