Genetic Variants of *Melanocortin 1 Receptor (MC1R)* Gene and Skin Cancer Risk Prediction

Shuvam Banerjee[1], Sandip Bhattacharjee[2], Pritha Bhattacharjee[3]

[1, 3]Environmental Epigenomics Lab, Department of Environmental Science, University of Calcutta

[2]Occupational Health, Health Management, Siemens Ltd, Gurgaon

Abstract: *The melanocortin-1 receptor (MC1R) gene (Mendelian Inheritance in Man 155555), plays an important role in the pigmentation process. The MC1R gene is low penetrating and highly polymorphic. A number of literature has studied the association of MC1R variants with skin cancer risk. Here, our objective is to shortlist the most potential MC1R variants based on sequence to structure and structure to function relationship. Out of total 1238 reported SNPs in MC1R, 258 were missense variants, of which, 11 were reported to be associated with cancer risk, according to earlier reports. Missense substitutions are of immense significance, as it causes the alterations in amino acid. Through in silico study, we have narrowed down three most potential variants i.e. R151C, R160W and D294H based on their functional prediction. Our next target was to identify other significant variations, if any, from remaining 247 SNPs, which are not reported earlier. We used several web based tools including SIFT, Mutation Taster, Polyphene2 and Meta-SNP and predicted six novel SNPs that might have crucial role in skin cancer. It is interesting to mention here that a number of literature identified the association of D84E (Aspartic acid to Glumatic acid at codon 84) and V60L (Valine to Leucine at codon 60) with skin cancer risk, although their "similarity" does not explain the association. We have attempted to explain the above "ambiguity" with secondary RNA structure of these two substitutions. To the best of our knowledge this is the first attempt to correlate the SNPs with structural prediction.*

Keywords: MC1R; Meta-SNP; Mutation Taster; Polyphene 2; Secondary Structure; SIFT

1. Introduction

Skin pigmentation is a human phenotype that varies among human population greatly and this variation is speculated to be adaptive[1]. Phenotypic variation in skin pigmentation of different human population has a correlation with latitude at the continental level [2]. Several studies presented significant differences in genetic variation between European, African and Asian population [2]. Melanocortin 1 receptor (*MC1R*) gene regulates human pigmentation and variation in the *MC1R* is associated with pigmentary phenotypes and risk of skin cancer types [3,4]. Thus *MC1R* plays a critical role in common skin cancers like squamous cell carcinoma, basal cell carcinoma and melanoma. *MC1R* is one of the low penetrating but important gene related to skin cancer. The *MC1R* gene locus is highly polymorphic in European population with more than 80 known variants [5] MC1R is a G protein–coupled receptor that binds to melanocortins, a class of pituitary peptide hormones. It includes adrenocorticotropic hormone (ACTH) and the different forms of melanocyte-stimulating hormone (MSH). The MC1R protein lies within the cell membrane, and melanocyte-stimulating hormone (MSH) released by the pituitary gland helps in its signaling. When it is activated by one of the variants of MSH, typically α-MSH, MC1R initiates a complex signaling cascade that leads to the production of the brown or black pigment named as eumelanin. In contrast, the receptor can also be antagonized by agouti signaling peptide (ASIP), which reverts the cell back to producing the yellow or red pigment known as phaeomelanin.

There are more than 100 non-synonymous SNP's that have been described [3,6].The recent database reports about 258 missense SNPs (https://www.ncbi.nlm.nih.gov/snp, last accessed April, 2017). Functional analysis of some of these variants showed partial loss in the ability of the receptor to stimulate the cAMP pathway, leading to a quantitative shift of melanin synthesis from eumelanin to phaeomelanin [3,7]. Red hair color is associated with phaeomelanin [8]. There is a good dependency of pigmentation characteristics with the *MC1R* variants. There are several polymorphisms in *MC1R* gene which are associated with the phenotype and are connected with melanoma and non-melanoma skin cancer risks [9]. Our first objective was to study whether all the missense SNPs reported to be associated with skin cancer risk, do correspond with effective functional alterations. Then, our next target was to predict novel SNPs based on in-silico studies that might have associated with skin cancer, although yet to be reported. Moreover, review of literature identified a number of variants which results in subtle changes in amino acids, to be associated with skin cancer risk. Attempt has been made to reanalyze the amino-acid substitutions in light of secondary structure prediction.

2. Methodology

2.1. SNP database and Listing of Missense substitutions: From NCBI, the missense variants of *MC1R* gene i.e. variants that results in amino acid alterations, were listed.

2.2. **Literature study and preliminary sorting of variants: A** thorough literature survey on "*MC1R* and skin cancer risk" was done based on global population data available. Variations were sorted based on their association with skin cancer risk, whether reported/not.

2.3. **Web-based shortlisting:** We used various web based application tools, like SIFT, Mutation Taster to predict whether the substitution of our variants of interest are "Tolerated/not Tolerated" and "disease causing". In

addition to this, whether this substitution ultimately lead to functional alteration and disease outcome, was predicted by another two web based tools, Polyphene2 and Meta SNP.

2.4. **Secondary structure generation of RNA sequences:** Secondary RNA structures were generated to understand how the subtle amino acid alterations might lead to disease outcome.

3. Results and Discussion

3.1. Results of SNP database and Literature survey

Out of total 1238 reported SNPs in *MC1R*, 258 were missense variants, of which, 11 were reported to be associated with cancer risk, according to literatures reported in Pubmed (Table-1).

3.2 Results of web-based shortlisting

- All 258 missense SNPs were tested using SIFT and Mutation taster, web-based tool application. SIFT predicts whether a single nucleotide alteration is tolerated or not to the sequence. Mutation Taster evaluates disease causing potential of sequence alterations [10]. Among 11 non synonymous SNPs (Table-1) which are reported to be associated with melanoma and NMSC in different populations, SIFT predicted 8 among them to be 'Not Tolerated'. Mutation taster predicts 7 of them to be 'Disease Causing'. The SNPs identified "risk allele" by either of the above mentioned tools, were sent for further analysis. The combined analysis of Polyphene2 and Meta SNP further shortlisted 3 SNPs, i.e. R151C, R160W and D294H (out of 8) as most potential variants, considering 0.85 as the threshold value (Fig. 1). The minimum and maximum score for both Polyphene2 and Meta SNP ranges from 0 and 1.

- In the similar way, remaining 247 missense SNPs were tested by both SIFT and Mutation taster. These identified 25 significant variants (Table-2). Analysis by Polyphene 2 and Meta-SNP, extracted 6 SNPs (out of 25) as most potential variants, considering 0.85 as the threshold value (Fig. 2).
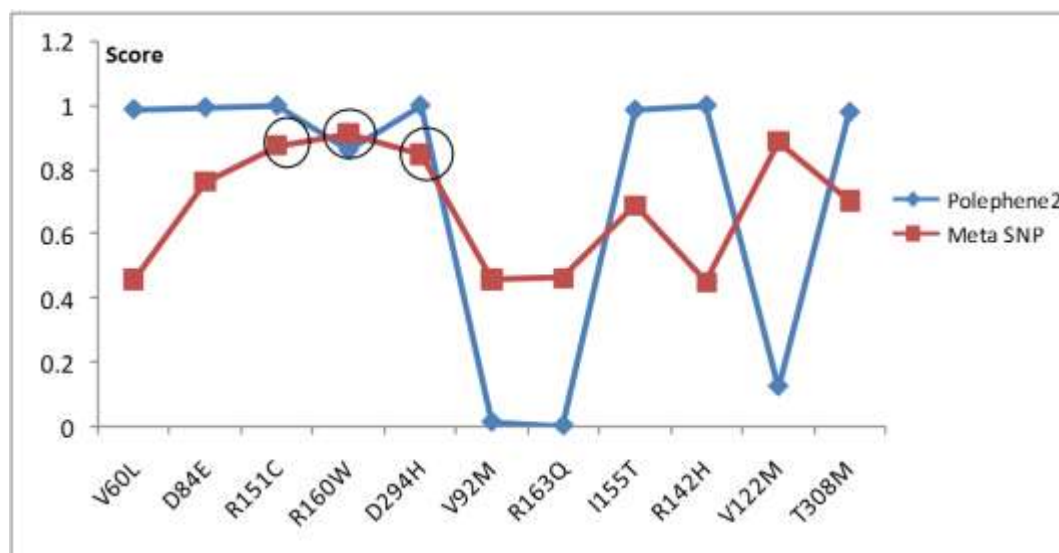

Fig 1: Polyphene2 and Meta SNP identified 3 most potential variants i.e. R151C, R160W and D294H out of 11 reported SNPs considering threshold value 0.85
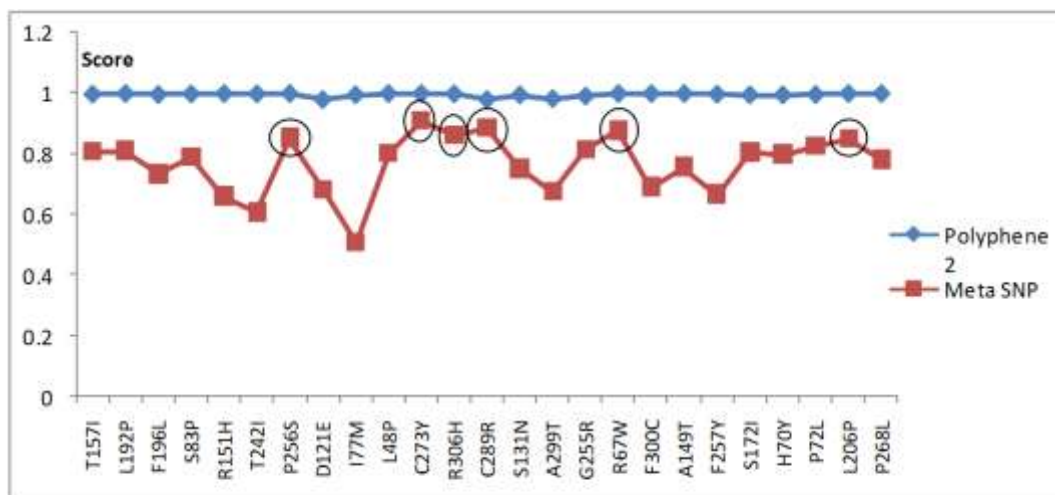
Fig.2: Polyphene2 and Meta SNP identified 6 novel variants i.e. P256S, C273Y, R306H, C289R, R67W and L206P (considering threshold value 0.85) out of 25 shortlisted SNPs based on SIFT and Mutation Taster

The other interesting observation was the association of Asp84Glu and Val60Leu with skin cancer risk. In Asp84Glu, Asp is coded by GAC/GAU, whereas Glu is coded by GAA/GAG. Therefore C>A or T>G change in nucleotide at third position, alters amino acid. Although both of these amino acids are non-essential for humans, and similar due to acidic as well as hydrophilic nature, it is very unlikely to find strong association of Asp84Glu variant in

Mc1R gene with skin cancer risk in different population (Table-3). Hence, we have attempted to study the secondary structure against each nucleotide change. It was found that no secondary structure alteration was there in case of GAC to GAA transversion (Fig 3A, 3B) but effective structural alteration could be observed only in case of GAC to GAG transversion (Fig 3A, 3C)
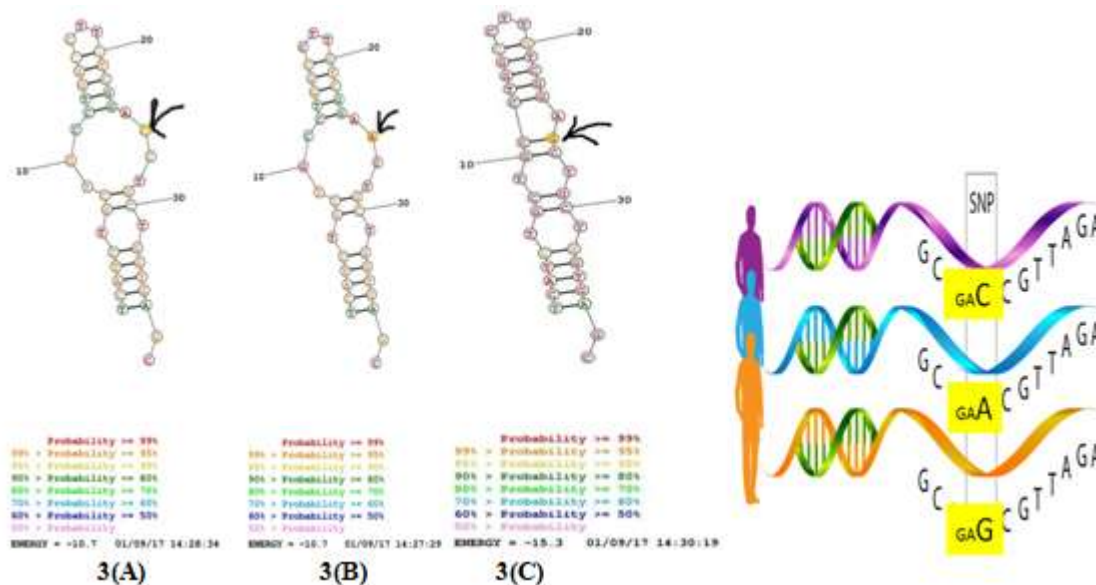

Figure 3: Secondary RNA Structures for GA**C**(3A) to GAA (3B) and GAG (3C) in **Asp84Glu**

It is indeed important to discuss about another substitution Val60Leu in *Mc1R* has significant contribution in skin cancer risk. Both are branched chain essential amino acids, as they are critical to human life and are particularly involved in stress, energy and muscle metabolism. Despite of their structural similarities, they have different metabolic routes and thus found to have significant association with disease risk. Valine can be coded by

GUU/GUC/GUA/GUG; whereas Leucine can be coded by CUU/CUC/CUA/CUG and UUA, UUG. Only G to T nucleotide change (GUG to UUG) was reported to be a significant SNP in database; however to the best of our knowledge we could not find any explanation in the literature for the same. We have generated secondary RNA structures for all possible substitution; and only GUG to UUG gave significant structural alterations (Fig. 4).
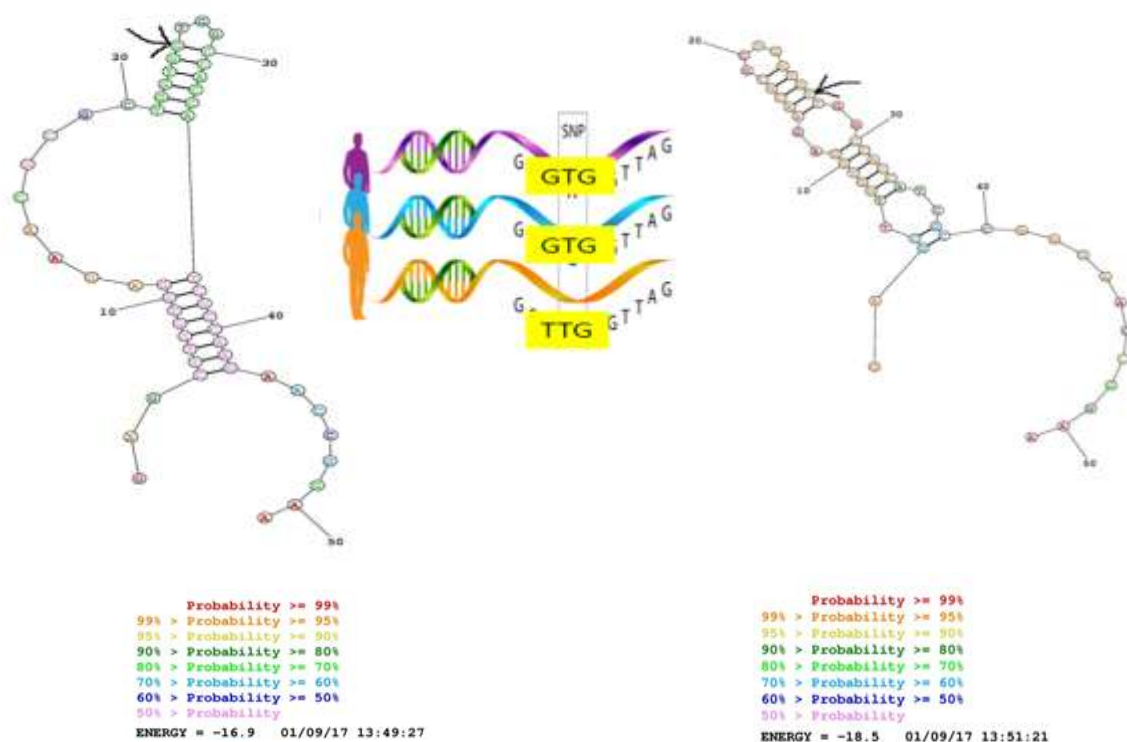
Figure 4: Secondary RNA Structures for GUG to UUG in Val60Leu

MC1R variants may have a role in carcinogenesis in addition to its influence on pigment variation. Experimental (*in vitro*) studies have shown that, besides its role in pigmentation, a-MSH, which binds to MC1R, is involved in other pathways, including anti-apoptotic DNA repair and anti-inflammatory pathways [11]. The genetic variantions in *MC1R* was studied in different population. Three most important SNPs were identified in our study as R151C, R160W and D294H. All these three variants showed association with skin cancer risk in different populations of all the continents except Africa and Asia.

The two most common non melanoma skin cancer (NMSC)s are basal cell carcinoma and squamous cell carcinoma. Binstock et al also suggested R151C and R160W variants are strongly associated with NMSC [12]. A systematic review and meta-analysis of *MC1R* variants estimate relative risk and population attributable fractions (PAF). PAF s ranged from 0.55% to 6.28% and the maximum PAF was observed for R151C [13]. Nan et al also found strongest association between R151C substitution and Basal Cell Carcinoma[14]. Genes, Environment and Melanoma (GEM) Study of Taylor et al and Kricker et al had also identified D294H as one of the important risk variant of *MC1R* gene in the populations of Australia, Canada, Italy and United States [4,15]. A bioluminescent assay also detected R151C, R160W and D294H changes with 'Red Hair Mutations' to be connected with melanoma and non-melanoma skin cancer comparatively higher than other *MC1R* polymorphisms [9].

Thus R151C, R160W and D294H are the most potential variants associated with skin cancer risk. Among 247 unreported SNP's of *MC1R* gene, P256S, C273Y, R306H, C289R, R67W and L206P are likely to have potential association with skin cancer risk. Further in-depth studies at population level is indeed important for validation. We have also explained the association of Asp84Glu and Val60Leu with skin cancer risk, despite of similarity of those amino acids. In both these cases, single nucleotide change, including C>G transversion for Asp84Glu and G>T transversion for Val60Leu showed secondary structure alteration and thus leading to considerable functional alteration of the protein. Therefore, we infer integrated knowledge of sequence-structure-function relationship is utmost important for SNP based association studies.

References

[1] Myles, S., Somel, M., Tang, K., Kelso, J., Stoneking, M., 2007. Identifying genes underlying skin pigmentation differences among human populations. Human Genetics. 120, 613-621

[2] Lao, O., de Gruijter,J.M., van Duijn, K., Navarro, A., Kayser,M., 2007. Signatures of positive selection in genes associated with Human skin pigmentation as revealed from analyses of Single Nucleotide Polymorphisms. Annals of Human Genetics. 71, 354-369

[3] Tagliabue, E., Fargnoli, MC., Gandini, S., Maisonneuve, P., Liu, F., Kayser, M., Nijsten, T., Han, J., Kumar, R., Gruis, NA., Ferrucci, L., Branicki, W., Dwyer, T., Blizzard, L., Helsing, P., Autier, P., García-Borrón, JC., Kanetsky, PA., Landi, MT., Little, J., Newton-Bishop, J., Sera, F., Raimondi, S., 2015. MC1R gene variants and non-melanoma skin cancer: a pooled-analysis from the M-SKIP project. Br J Cancer. 113, 354-363

[4] Taylor, NJ., Busam, KJ., From, L., Groben, PA., Anton-Culver, H., Cust, AE., Begg, CB., Dwyer, T., Gallagher, RP., Gruber, SB., Orlow, I., Rosso, S., Thomas,

NE., Zanetti, R., Rebbeck, TR., Berwick, M., Kanetsky, PA., 2015. Inherited Variation at MC1R and Histological Characteristics of Primary Melanoma. PLoS One, 10.

[5] Pasquali, E., García-Borrón, JC., Fargnoli, MC., Gandini, S., Maisonneuve, P., Bagnardi, V., Specchia, C., Liu, F., Kayser, M., Nijsten, T., Nagore, E., Kumar, R., Hansson, J., Kanetsky, PA., Ghiorzo, P., Debniak, T.,Branicki, W., Gruis, NA., Han, J., Dwyer, T., Blizzard, L., Landi, MT., Palmieri, G., Ribas, G., Stratigos, A., Council, ML., Autier, P., Little, J., Newton-Bishop, J., Sera, F.,Raimondi, S., 2014. MC1R variants increased the risk of sporadic cutaneous melanoma in darker-pigmented Caucasians: a pooled-analysis from the M-SKIP project. Int J Cancer. 136, 618-631

[6] Pérez Oliva, AB., Fernéndez, LP., Detorre, C., Herráiz, C., Martínez-Escribano, JA., Benítez, J., Lozano Teruel, JA., García-Borrón, JC., Jiménez-Cervantes, C., Ribas, G., 2009. Identification and functional analysis of novel variants of the human melanocortin 1 receptor found in melanoma patients. Hum Mutat. 30, 811-822.

[7] Duffy, DL., Box, NF., Chen, W., Palmer, JS., Montgomery, GW., James, MR., Hayward, NK., Martin, NG., Sturm, RA., 2004. Interactive effects of MC1R and OCA2 on melanoma risk phenotypes. Hum Mol Genet. 13, 447-461

[8] Box, NF., Wyeth, JR., O'Gorman, LE., Martin, NG., Sturm, RA., 1997. Characterization of melanocyte stimulating hormone receptor variant alleles in twins with red hair. Hum Mol Genet. 6, 1891-1897

[9] Bashmakova, EE., Krasitskaya, VV., Bondar, AA., Kozlova, AV., Ruksha, TG., Frank, LA., 2015. Bioluminescent assay to detect melanocortin-1 receptor (MC1R) polymorphisms (R160W, R151C, and D294H). Mol Biol (Mosk). 49,953-958.

[10] Schwarz, JM., Rödelsperger, C., Schuelke, M., Seelow, D., 2010. Mutation Taster evaluates disease causing potential of sequence alterations. Nature Methods. 7, 575-576.

[11] Viki, B.S., Zalfa, A., 2016. Significance of the Melanocortin 1 and Endothelin B Receptors in Melanocyte Homeostasis and Prevention of Sun-Induced Genotoxicity. Front Genet. 7, 146.

[12] Binstock, M., Hafeez, F.,Metchnikoff, C., Arron, ST., 2014. Single-nucleotide polymorphisms in pigment genes and nonmelanoma skin cancer predisposition: a systematic review. Br J Dermatol. 171,713-721.

[13] Williams, PF., Olsen, CM., Hayward, NK., Whiteman, DC., 2011. Melanocortin 1 receptor and risk of cutaneous melanoma: a meta-analysis and estimates of population burden. Int J Cancer. 129, 1730-1740.

[14] Nan, H., Xu, M., Kraft, P., Qureshi, AA., Chen, C., Guo, Q., Hu, FB., Curhan, G., Amos, CI., Wang, LE,, Lee, JE., Wei, Q., Hunter, DJ., Han, J.,2011. Genome-wide association study identifies novel alleles associated with risk of cutaneous basal cell carcinoma and squamous cell carcinoma. Hum Mol Genet. 20, 3718-3724.

[15] Kricker, A., Armstrong, BK., Goumas, C., Kanetsky, P., Gallagher, RP., Begg, CB., Millikan, RC., Dwyer, T., Rosso, S., Marrett, LD., Thomas, NE., Berwick, M., 2010. MC1R genotype may modify the effect of sun exposure on melanoma risk in the GEM study. Cancer Causes Control. 21, 2137-2147.

[16] Puig-Butillé, JA., Carrera, C., Kumar, R., Garcia-Casado, Z., Badenas, C., Aguilera, P., Malvehy, J., Nagore, E., Puig, S., 2013. Distribution of *MC1R* variants among melanoma subtypes: p.R163Q is associated with Lentigo Maligna Melanoma in a Mediterranean population. Br J Dermatol. 169, 804-811.

[17] Helsinq,P., Nymoen, DA., Rootwelt, H., Vårdal, M., Akslen, LA., Molven, A., Andresen, PA.,2012. MC1R, ASIP, TYR, and TYRP1 gene variants in a population-based series of multiple primary melanomas. Genes Chromosomes Cancer. 51, 654-661

[18] Cust, AE., Goumas, C., Holland, EA., Agha-Hamilton, C., Aitken, JF., Armstrong, BK., Giles, GG., Kefford, RF., Schmid, H., Hopper, JL., Mann, GJ., Jenkins, MA., 2012. MC1R genotypes and risk of melanoma before age 40 years: a population-based case-control- family study. Int J Cancer. 131, 269-281

[19] Avilés, JA., Lázaro, P., Fernández, LP., Benítez, J., Ibarrola-Villava, M., Ribas G.,2012. Phenotypic and histologic characteristics of cutaneous melanoma in patients with melanocortin-1 receptor polymorphisms. Actas Dermosifiliogr. 103,44-50.

[20] Demenais, F., Mohamdi, H., Chaudru, A., Goldstein, M., Newton Bishop, J. A., Bishop, D.T., Kanetsky, P.A., Hayward, N.K., Gillanders, E., Elder, D.E., Avril, M.F., Azizi, E., van Belle, P., Bergman, W., Bianchi-Scarrà, G., Bressac-de Paillerets, B., Calista, D., Carrera, C., Hansson, J., Harland, M., Hogg, D., Höiom, V., Holland, E.A., Ingvar, C., Landi, M.T., Lang, J.M., Mackie, R.M., Mann, G.J., Ming, M.E., Njauw, C.J., Olsson, H., Palmer, J., Pastorino, L., Puig, S., Randerson-Moor, J., Stark, M., Tsao, H., Tucker, M.A., van der Velden, P., Yang, X.R., Gruis, N., 2010. J Natl Cancer Inst. 102, 1568-1583

Table 1: Reported *MC1R* variants from database and risk prediction using SIFT, Mutation Taster, Polyphene 2 and Meta SNP

Variant ID	Residue change	SIFT Predictions	Mutation Taster	Poly Phene2 Results	Meta SNP Predictions	Literature availability (PMID no)*
rs1805005	Val60Leu	Tolerated	Polymorphism	0.988 (Pr_D)	0.460 Neutral	26103569, 24917043, 23647022, 22464597, 21737053, 21128237, 20876876
rs1805006	Asp84Glu	Not Tolerated	Disease Causing	0.994 (Pr_D)	0.764; Disease	26103569,25790105, 24917043, 23647022,22447455, 22095472, 20721616
rs1805007	Arg151Cys	Not Tolerated	Disease Causing	1.000 (Pr_D)	0.878; Disease	26710775, 26103569, 25790105, 25319428, 24170137, 23647022, 23522749, 22447455, 22095472, 21700618, 21128237, 20876876, 20721616
rs1805008	Arg160Trp	Not Tolerated	Polymorphism	0.861 (Po_D)	0.913;	26710775, 26103569, 25790105, 25319428,

					Disease	23647022, 22497519, 22464597, 22095472, 21737053, 20876876, 20721616, 20629734
rs1805009	Asp294His	Not Tolerated	Disease Causing	1.000 (Pr_D)	0.850; Disease	26710775, 26103569, 25790105, 23647022, 22464597, 22095472, 21737053, 20721616, 20629734
rs2228479	Val92Met	Tolerated	Disease Causing	0.015 (Benign)	0.461; Neutral	26103569, 23647022, 22464597, 21737053, 20876876, 20629734
rs885479	Arg163Gln	Tolerated	Polymorphism	0.004 (Benign)	0.465; Neutral	26103569, 24660985, 24170137, 23647022, 22464597, 21737053
rs1110400	Ile155Thr	Not Tolerated	Disease Causing	0.986 (Pr_D)	0.690; Disease	23647022, 22464597, 22095472, 21737053, 21128237, 20629734
rs11547464	Arg142His	Not Tolerated	Disease Causing	1.000 (Pr_D)	0.453; Neutral	25790105, 23647022, 22095472,21128237
rs201192930	Val122Met	Tolerated	Polymorphism	0.126 (Benign)	0.889; Disease	23647022
rs375127718	Thr308Met	Not Tolerated	Disease Causing	0.979 (Pr_D)	0.704; Disease	20629734

* PUBMED ids are given as references here (Table-1)

Table 2: SIFT and Mutation taster identified <u>novel</u> 25 *MC1R* variants and their risk prediction using Polyphene 2 and Meta SNP in the present study

Variant ID	Residue change	Effectivity in functional alteration	Poly Phene-2 Results	Meta SNP Predictions
rs104894524	Thr157Ile	Minor	0.998 (Pr_D)	0.807; Disease
rs587783375	Leu192Pro	Minor	1.000 (Pr_D)	0.810; Disease
rs3212366	Phe196Leu	Average	0.997 (Pr_D)	0.734; Disease
rs34474212	Ser83Pro	Minor	0.999 (Pr_D)	0.791; Disease
rs149922657	Arg151His	Minor	1.000 (Pr_D)	0.661; Disease
rs200051702	Thr242Ile	Minor	0.999 (Pr_D)	0.608; Disease
rs200215218	Pro256Ser	Minor	1.000 (Pr_D)	0.852; Disease
rs200616835	Asp121Glu	Minor	0.981 (Pr_D)	0.683; Disease
rs200759505	Ile77Met	Minor	0.996 (Pr_D)	0.509; Disease
rs201787533	Leu48Pro	Average	1.000 (Pr_D)	0.803; Disease
rs368281517	Cys273Tyr	Major	1.000 (Pr_D)	0.909; Disease
rs368507952	Arg306His	Minor	0.999 (Pr_D)	0.862; Disease
rs369542041	Cys289Arg	Major	0.981(Pr_D)	0.886; Disease
rs370094672	Ser131Asn	Major	0.996 (Pr_D)	0.751; Disease
rs370472871	Ala299Thr	Minor	0.984 (Pr_D)	0.675; Disease
rs371214731	Gly255Arg	Major	0.992 (Pr_D)	0.815; Disease
rs372590533	Arg67Trp	Major	1.000 (Pr_D)	0.878; Disease
rs373872609	Phe300Cys	Major	1.000 (Pr_D)	0.691; Disease
rs374423188	Ala149Thr	Minor	1.000 (Pr_D)	0.756; Disease
rs376508354	Phe257Tyr	Minor	0.998 (Pr_D)	0.667; Disease
rs376670171	Ser172Ile	Minor	0.996 (Pr_D)	0.805; Disease
rs377122753	His70Tyr	Major	0.996 (Pr_D)	0.798; Disease
rs377297107	Pro72Leu	Average	0.997 (Pr_D)	0.826; Disease
rs377499038	Leu206Pro	Average	1.000 (Pr_D)	0.850; Disease

Table 3: *MC1R* Asp84Glu and association with skin cancer risk reported in global population

Type of study	No of Case	No of Control	Population	Odds ratio	Reference
Pooled Analysis	3527	9391	European	2.66 (1.06-6.65)	Taqliabue et al., 2015
GEM Study	2424	N/A	Australia, Canada, Italy, United States	2.39 (1.40-4.09)	Taylor et al., 2015
Pooled Analysis	5160	12119	Darker-pigmented Caucasians	2.74 (1.53-4.89)	Pasquali et al., 2015
Association study	1679	N/A	Mediterranean	1%	Puiq-Butille et al., 2013[16]
Association study	388	420	Norwegian	5.77 (1.97-16.90)	Helsinq et al., 2012[17]
Family based Association Study	565	927	Australian	2.08 (0.91–4.72)	Cust et al., 2012[18]
GEM Study	1018	1875	Australia, Canada, Italy and the USA	1.39 (1.05-1.84)	Kricker et al., 2010

Table 4: *MC1R* Val 60 Leu and association with skin cancer risk reported in global population

Type of study	No of Case	No of Control	Population	Odds ratio/ Percentage	Reference:
Pooled Analysis	3527	9591	European	1.42 (1.19-1.70)	Taqliabue et al., 2015
Pooled Analysis	5160	12119	Darker-pigmented Caucasians	1.47 (1.17-1.84)	Pasquali et al., 2015
Association study	1679	N/A	Mediterranean	30%	Puiq-Butille et al., 2013
Association study	224		Spanish	29.9%	Aviles et al., 2012[19]
Meta-Analysis	5164	45066	14 different population	1.18 (1.04-1.35)	Williams et al., 2011
Family based Association Study	473	342	European, North American, and Australian	3.42 (2.10 to 5.58)	Demenais et al., 2010[20]