SNPs, linkage disequilibrium, and chronic mountain sickness in Tibetan Chinese

Norman E Buroker1
Xue-Han Ning1,2,†
Zhao-Nian Zhou3
Kui Li4
Wei-Jun Cen4
Xiu-Feng Wu3
Wei-Zhong Zhu5
C Ronald Scott1
Shi-Han Chen1

1Department of Pediatrics, University of Washington, 2Division of Cardiology, Seattle Children’s Hospital Research Foundation, Seattle, WA, USA; 3Laboratory of Hypoxia Physiology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai, China; 4Lhasa People Hospital, Lhasa, Tibet; 5Center for Cardiovascular Biology and Regenerative Medicine, University of Washington, Seattle, WA, USA

†Xue-Han Ning passed away on April 20, 2015

Abstract: Chronic mountain sickness (CMS) is estimated at 1.2% in Tibetans living at the Qinghai–Tibetan Plateau. Eighteen single-nucleotide polymorphisms (SNPs) from nine nuclear genes that have an association with CMS in Tibetans have been analyzed by using pairwise linkage disequilibrium (LD). The SNPs included are the angiotensin-converting enzyme (rs4340), the angiotensinogen (rs699), and the angiotensin II type 1 receptor (AGTR1) (rs5186) from the renin–angiotensin system. A low-density lipoprotein apolipoprotein B (rs693) SNP was also included. From the hypoxia-inducible factor oxygen signaling pathway, the endothelial Per-Arnt-Sim domain protein 1 (ECPAS1) and the egl nine homolog 1 (ENGL1) (rs480902) SNPs were included in the study. SNPs from the vascular endothelial growth factor (VEGF) signaling pathway included are the v-akt murine thymoma viral oncogene homolog 3 (rs4590656 and rs2291409), the endothelial cell nitric oxide synthase 3 (rs1007311 and rs1799983), and the (VEGFA) (rs699947, rs34357231, rs79469752, rs13207351, rs28357093, rs1570360, rs2010963, and rs3025039). An increase in LD occurred in 40 pairwise comparisons, whereas a decrease in LD was found in 55 pairwise comparisons between the controls and CMS patients. These changes were found to occur within and between signaling pathways, which suggests that there is an interaction between SNP alleles from different areas of the genome that affect CMS.

Keywords: HAS, Qinghai-Tibetan Plateau, RAS, HIF, VEGF, pathways, miRNA

Introduction

Chronic mountain sickness (CMS) is a syndrome involving the loss of adaptation to life at high altitude, which occurs in some residents residing in this environment.1,2 Older age in these permanent residents has been linked to the sickness.3 The susceptibility to CMS between and within human populations at high altitudes has been shown to have great variability.4 The precise pathogenesis of CMS is not well understood, but the lack of oxygen at high altitudes is considered to be the major factor.5–8 Which raises the question of why some individuals are susceptible to the sickness while others are not. At the same altitude, the prevalence of the disease is estimated to be 15% in Quechua Andeans and 1.2% in Tibetans.9 Based on the current genetic data, it has been suggested that Tibetans may be one of the oldest high-altitude adapted ethnic groups in the world with origins from the Neolithic period.10–13 Which might explain the lower occurrence of the sickness in Tibetans.

Genome-wide association studies (GWAS) have identified >15 thousand single-nucleotide polymorphisms (SNPs) associated with various diseases and traits.14,15 Some of these SNPs have been examined in association with high-altitude sickness (HAS) among Han and Tibetan Chinese at the Qinghai–Tibetan Plateau in China.16–20 Recently,
the association of SNPs that create potential changes in transcriptional factor binding sites (TFBS) and HAS have been examined.\textsuperscript{20,21} To this end, punitive changes in TFBS have been reported to be created by 17 SNPs among nine genes significantly associated with acute mountain sickness (AMS) in Han Chinese.\textsuperscript{20} In this study, linkage disequilibrium (LD) was used to analyze intra- and intergenic SNP alleles and alterations in TFBS. Pairwise LD was computed between the SNP alleles within a control group and an AMS group. An increase in LD was found to occur in 32 pairwise comparisons, whereas a decrease in LD was found in 22 comparisons between the groups.\textsuperscript{20} The pairwise LD changes were found within and between genes involving signaling pathways and systems indicating the interaction of SNP alleles or punitive TFBS from different areas of the genome. The current study examined the LD between 18 SNPs from the same nine genes in CMS Tibetan patients at the Qinghai–Tibetan Plateau in China since these SNPs have previously been associated with HAS.\textsuperscript{16–19}

**Materials and methods**

**Study groups**

The Tibetan CMS patients sampled in the study had been hospitalized and diagnosed by physicians at the Lhasa People Hospital (Tibet, China, at 3,670 m above sea level).\textsuperscript{17–19} Healthy Tibetans were randomly selected from the Lhasa Hospital (Tibet, China, at 3,670 m above sea level).\textsuperscript{17–19} In this study, linkage disequilibrium (LD) has been reported to be created by 17 SNPs among nine genes significantly associated with acute mountain sickness (AMS) in Han Chinese.\textsuperscript{20} In this study, linkage disequilibrium (LD) was used to analyze intra- and intergenic SNP alleles and alterations in TFBS. Pairwise LD was computed between the SNP alleles within a control group and an AMS group. An increase in LD was found to occur in 32 pairwise comparisons, whereas a decrease in LD was found in 22 comparisons between the groups.\textsuperscript{20} The pairwise LD changes were found within and between genes involving signaling pathways and systems indicating the interaction of SNP alleles or punitive TFBS from different areas of the genome. The current study examined the LD between 18 SNPs from the same nine genes in CMS Tibetan patients at the Qinghai–Tibetan Plateau in China since these SNPs have previously been associated with HAS.\textsuperscript{16–19}

**Genotyping**

The methods for genotyping each SNP have been previously outlined and reported.\textsuperscript{16–19} The genes and their SNPs used in the present study are summarized in Table 1.

**Statistical analysis**

The statistical methods used to genotype samples have been previously discussed.\textsuperscript{16–19} Pairwise LD\textsuperscript{22} was computed between all SNPs from the Tibetan CMS and control groups area to serve as control subjects. The average age of CMS patients was 53.6 years, whereas that of the non-CMS Tibetan controls was 30.5 years. All Tibetans sampled in the study signed an informed consent approved by the Human Ethics Committee of the Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences. The study was approved by the Shanghai Institutes for Biological Sciences review board.

**Sampling**

Buccal brush samples were collected at the Lhasa People Hospital from 45 Tibetans diagnosed with CMS during the high occurrence period (spring and winter). Control samples consisting of 34 Tibetan unaffected natives were also collected via buccal brush. The controls were found to be in good health upon physical examination by doctors at the Lhasa People Hospital.\textsuperscript{16–19}

**Table 1 Genes and their SNPs that have been found to be associated with chronic mountain sickness**

<table>
<thead>
<tr>
<th>Protein and gene symbol</th>
<th>Chromosome</th>
<th>SNP</th>
<th>SNP location</th>
<th>Mutation</th>
<th>LD identity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Angiotensin I-converting enzyme (ACE)</td>
<td>17q23.3</td>
<td>rs4340</td>
<td>Intron 16</td>
<td>288bp Indel/(ALU)</td>
<td>ACE</td>
</tr>
<tr>
<td>Angiotensinogen (AGT)</td>
<td>1q43.2</td>
<td>rs699</td>
<td>Exon 1</td>
<td>c.803T&gt;C, p. Met268Thr</td>
<td>AGT</td>
</tr>
<tr>
<td>Angiotensin II receptor, type 1 (AGTR1)</td>
<td>3q24</td>
<td>rs5186</td>
<td>3'UTR</td>
<td>c.*86A&gt;G</td>
<td>AGTR1</td>
</tr>
<tr>
<td>v-Akt murine thymoma viral oncogene homolog 3 (AKT3)</td>
<td>1q44</td>
<td>rs4590636</td>
<td>Intron 1</td>
<td>c.461-3654C&gt;T</td>
<td>AKT3-45</td>
</tr>
<tr>
<td>v-Akt murine thymoma viral oncogene homolog 3 (AKT3)</td>
<td>1q44</td>
<td>rs2291409</td>
<td>Intron 8</td>
<td>c.819+4031G&gt;A</td>
<td>AKT3-22</td>
</tr>
<tr>
<td>Apolipoprotein B (APOB)</td>
<td>2p24.1</td>
<td>rs693</td>
<td>Exon 26</td>
<td>c.7545C&gt;T, p.Thr2515Thr</td>
<td>APOB</td>
</tr>
<tr>
<td>Egf-9 family hypoxia-inducible factor 1 (EGLN1)</td>
<td>1q42.2</td>
<td>rs480902</td>
<td>Intron 1</td>
<td>c.892-21782A&gt;T</td>
<td>EGLN1-48</td>
</tr>
<tr>
<td>Endothelial PAS domain protein 1 (EPAS1)</td>
<td>2p21</td>
<td>unknown</td>
<td>Intron 5</td>
<td>c.657-5C&gt;G</td>
<td>EPAS1</td>
</tr>
<tr>
<td>Nitric oxide synthase 3 (eNOS3)</td>
<td>7q36.1</td>
<td>rs1007311</td>
<td>Intron 6</td>
<td>c.817-26A&gt;G</td>
<td>eNOS3-10</td>
</tr>
<tr>
<td>Nitric oxide synthase 3 (eNOS3)</td>
<td>7q36.1</td>
<td>rs1799983</td>
<td>Exon 7</td>
<td>c.894-1154T&gt;C</td>
<td>eNOS3-17</td>
</tr>
<tr>
<td>Vascular endothelial growth factor A (VEGFA)</td>
<td>6p21.1</td>
<td>rs699947</td>
<td>(-)2576 TSS</td>
<td>c.-2576C&gt;A</td>
<td>VEGFA-69</td>
</tr>
<tr>
<td>Vascular endothelial growth factor A (VEGFA)</td>
<td>6p21.1</td>
<td>rs34357231</td>
<td>(-)2550 TSS</td>
<td>c.-2550-2568D&gt;1</td>
<td>VEGFA-34</td>
</tr>
<tr>
<td>Vascular endothelial growth factor A (VEGFA)</td>
<td>6p21.1</td>
<td>rs79469752</td>
<td>(-)1203 TSS</td>
<td>c.-663C&gt;T</td>
<td>VEGFA-79</td>
</tr>
<tr>
<td>Vascular endothelial growth factor A (VEGFA)</td>
<td>6p21.1</td>
<td>rs13207351</td>
<td>(-)1190 TSS</td>
<td>c.-650A&gt;G</td>
<td>VEGFA-13</td>
</tr>
<tr>
<td>Vascular endothelial growth factor A (VEGFA)</td>
<td>6p21.1</td>
<td>rs28357093</td>
<td>(-)1179 TSS</td>
<td>c.-639A&gt;C</td>
<td>VEGFA-28</td>
</tr>
<tr>
<td>Vascular endothelial growth factor A (VEGFA)</td>
<td>6p21.1</td>
<td>rs1570360</td>
<td>(-)1154 TSS</td>
<td>c.-614A&gt;G</td>
<td>VEGFA-15</td>
</tr>
<tr>
<td>Vascular endothelial growth factor A (VEGFA)</td>
<td>6p21.1</td>
<td>rs2010963</td>
<td>(-)634 TSS</td>
<td>c.-634C&gt;G</td>
<td>VEGFA-20</td>
</tr>
<tr>
<td>Vascular endothelial growth factor A (VEGFA)</td>
<td>6p21.1</td>
<td>rs3025039</td>
<td>3'UTR</td>
<td>c.-237C&gt;T</td>
<td>VEGFA-30</td>
</tr>
</tbody>
</table>

**Notes:** Location of gene chromosome, SNP location in the gene, and the resulting genetic mutation as well as the LD identity using in the analysis are listed.

**Abbreviations:** LD, linkage disequilibrium; SNP, single-nucleotide polymorphism; TSS, transcriptional start site.
The degree of genetic linkage between the 18 SNPs in each study group was estimated as Lewontin’s coefficient $D'$, where no color ($D'=0$) indicates that LD is weak or nonexistent and the dark red ($D'=1$) indicates that there exists strong pairwise LD between SNPs (Figure 1). The change in LD between the control and CMS groups among all SNPs is presented in Table 2.

### miRNA analysis
The miRdSNP$^{23}$ and miRBase$^{24-28}$ databases were used to identify the has-miR-591 miRNA-predicted target site in association with the VEGFA (rs3025039) SNP in the 3’UTR of the $VEGFA$ gene.

### Results
Pairwise LD estimates were obtained between 18 SNPs from nine genes that had been previously found to be associated with CMS.$^{17-19}$ The SNPs occur in the renin–angiotensin system (RAS) (ACE, AGT, and AGTR1), GNB3, LDL, APOB, the HIF oxygen signaling pathway EPAS1 and EGLN1), and the vascular endothelial growth factor (VEGF) signaling pathway (AKT3, eNOS3, and VEGFA) genes (Tables 1 and 2; Figures 1 and 2). In the pairwise estimates, the ACE with AKT3-45, APOB, VEGFA-13, and VEGFA-15 SNP pairs exhibit an increase in LD for the CMS study group compared with the control group and a decrease in LD in ACE with AGT, eNOS3-10, eNOS3-17, VEGFA-79, VEGFA-28, and VEGFA-30 pairs (Figure 1 and Table 2). Also, in the pairwise estimates, the AGT with VEGFA-69 and VEGFA-34 pairs exhibit an increase in LD for the CMS study group compared with the control group and a decrease in LD in AGT with EGLN1-48, eNOS3-10, VEGFA-79, VEGFA-13, VEGFA-15, and VEGFA-20 pairs (Figure 1; Table 2). In addition, in these estimates, the AGTR1 with APOB, EGLN1-48, EPAS1, VEGFA-69, VEGFA-34, VEGFA-13, and VEGFA-15 pairs exhibit an increase in LD for the CMS study group compared with the control group and a decrease in LD in AGTR1 with AKT3-45, AKT3-22, eNOS3-17, and VEGFA-28 pairs (Figure 1; Table 2). Similar results can be seen for the other pairwise LD estimates for all the SNPs (Figure 1; Table 2) including the association of SNPs between these signaling pathways and systems (Figure 2).

The SNPs of the AKT3, eNOS3, and VEGFA genes in the VEGF signaling pathway (a growth factor activator for angiogenesis) and the SNPs of the EGLN1 and EPAS1 genes in the HIF oxygen signaling pathway exhibit an increase in pairwise LD for the CMS group compared with the control Tibetan Chinese groups (Figure 1; Table 2) both within and between the pathways (Figure 2). Besides the SNP (rs4340) ACE-I allele, there are only two other SNPs (rs4590656 and rs3025039) AKT3-C and VEGFA-C allele, respectively, that generate potential unique hypoxia transcriptional factor binding (HIF1a:ARNT)$^{20}$.

A decrease in LD between SNPs for the CMS group compared with the Tibetan Chinese non-CMS group could indicate that some genes are responding to different attributes of the CMS sickness. As an example, the $VEGFA-30$ SNP (rs3025039) in the 3’UTR exhibits a tremendous decrease in LD in the CMS group compared with the non-CMS group with respect to the ACE, EGLN1-48, EPAS1, eNOS3-10, eNOS3-17, VEGFA-69, VEGFA-34, VEGFA-79, and VEGFA-20 SNPs (Figure 1; Table 2). Apparently, when CMS occurs among Tibetan individuals, there is less interaction between the $VEGFA-30$ SNP (rs3025039) and these other SNPs. A
The scan of the miRdSNP and miRBase databases revealed that the \( VEGFA \) (rs3025039) SNP alters the miRNA (has-miR-591)-predicted target site CA\[C/T\]GGTC in the 3'UTR of the gene. This SNP has been associated with several human diseases (see “Discussion” section) as well as CMS. Overall, there were 40 pairwise increases and 55 pairwise decreases in the gene.

### Table 2
Changes in pairwise LD from Figure 1 between 18 SNPs in the CMS Tibetan Chinese group compared with the non-CMS Tibetan Chinese group

<table>
<thead>
<tr>
<th>Increase in LD</th>
<th>Decrease in LD</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 ACE vs AKT3-45</td>
<td>21 APOB vs eNOS3-17</td>
</tr>
<tr>
<td>2 ACE vs APOB</td>
<td>22 APOB vs VEGFA-69</td>
</tr>
<tr>
<td>3 ACE vs VEGFA-13</td>
<td>23 APOB vs VEGFA-34</td>
</tr>
<tr>
<td>4 ACE vs VEGFA-15</td>
<td>24 APOB vs VEGFA-13</td>
</tr>
<tr>
<td>5 AGT vs VEGFA-69</td>
<td>25 APOB vs VEGFA-15</td>
</tr>
<tr>
<td>6 AGT vs VEGFA-34</td>
<td>26 APOB vs VEGFA-30</td>
</tr>
<tr>
<td>7 AGTR1 vs APOB</td>
<td>27 EGLN1-48 vs VEGFA-28</td>
</tr>
<tr>
<td>8 AGTR1 vs EGLN1-48</td>
<td>28 EGLN1-48 vs VEGFA-20</td>
</tr>
<tr>
<td>9 AGTR1 vs EPAS1</td>
<td>29 EPAS1 vs VEGFA-15</td>
</tr>
<tr>
<td>10 AGTR1 vs VEGFA-69</td>
<td>30 eNOS3-10 vs eNOS3-17</td>
</tr>
<tr>
<td>11 AGTR1 vs VEGFA-34</td>
<td>31 VEGFA-69 vs VEGFA-13</td>
</tr>
<tr>
<td>12 AGTR1 vs VEGFA-13</td>
<td>32 VEGFA-69 vs VEGFA-15</td>
</tr>
<tr>
<td>13 AGTR1 vs VEGFA-15</td>
<td>33 VEGFA-69 vs VEGFA-20</td>
</tr>
<tr>
<td>14 AKT3-45 vs APOB</td>
<td>34 VEGFA-34 vs VEGFA-13</td>
</tr>
<tr>
<td>15 AKT3-22 vs APOB</td>
<td>35 VEGFA-34 vs VEGFA-20</td>
</tr>
<tr>
<td>16 AKT3-22 vs eNOS3-10</td>
<td>36 VEGFA-34 vs VEGFA-20</td>
</tr>
<tr>
<td>17 AKT3-22 vs eNOS3-17</td>
<td>37 VEGFA-79 vs VEGFA-13</td>
</tr>
<tr>
<td>18 AKT3-22 vs VEGFA-30</td>
<td>38 VEGFA-79 vs VEGFA-28</td>
</tr>
<tr>
<td>19 APOB vs EGLN1-48</td>
<td>39 VEGFA-79 vs VEGFA-20</td>
</tr>
<tr>
<td>20 APOB vs EPAS1</td>
<td>40 VEGFA-13 vs VEGFA-20</td>
</tr>
</tbody>
</table>

Note: Increase and decrease in pairwise LD in the CMS group compared with the Tibetan Chinese control group are listed.

Abbreviations: LD, linkage disequilibrium; SNP, single-nucleotide polymorphism; CMS, chronic mountain sickness.

**Figure 2** Pairwise LD for 18 SNPs in the CMS Tibetan Chinese group compared with the non-CMS Tibetan Chinese groups from Table 2 is graphed between the pathways. Figure (A) represents an increase in LD in the CMS group compared with the control group, whereas figure (B) represents a decrease in LD. The numbers are the pairwise associations.

Abbreviations: APOB, apolipoprotein B; HIF, hypoxia-inducible factor oxygen signaling pathway; RAS, renin–angiotension system; VEGFA, vascular endothelial growth factor signaling pathway; LD, linkage disequilibrium; SNP, single-nucleotide polymorphism; CMS, chronic mountain sickness.
LD in the CMS group compared with the non-CMS control group (Figure 1; Table 2).

Discussion

Of the GWAS-identified disease or trait-predisposing SNPs, only 7% of these are located in protein-coding regions of the genome, and the remaining 93% are located within noncoding areas, such as regulatory or intergenic regions. SNPs occurring in the regulatory region of a gene are drawing more attention, where a single nucleotide change can alter the binding site of a transcriptional factor (TF), which can result in multiple consequences. In most of the cases, the SNP will neither change the TFBS interaction nor alter gene expression since a TF can usually recognize a number of different binding sites in various areas of the gene. However, in other cases, the SNP may increase or decrease the ability of TF to bind DNA which would result in allele-specific gene expression. Also, an SNP may eliminate the TFBS or generate a new binding site for another TF. In such cases, the gene would no longer be regulated by the original TF. Therefore, these functional regulatory SNPs that occur in TFBS may cause changes in gene expression, phenotypes, and susceptibility to environmental exposure. Numerous examples and several reviews have been written about rSNPs associated with disease susceptibility. Such rSNPs occurring in the regulatory regions of genes have been associated with disease or sickness. These regulatory regions host the binding sites for transcription factors involved in gene expression.

LD is defined as the non-random association between the alleles of two or more loci in a given population. LD between SNP alleles in the regulatory region of a gene has been used as a method of identifying haplotype associations relating to sickness or disease in a population. This is achieved when the levels of LD between SNP alleles within haplotypes are seen to change substantially in a sickness group when compared with a baseline population. In such cases, LD between SNP alleles can be used to identify punitive binding changes for TFs responsible for gene regulation. Such TFBS changes may result in human disease or sickness. In this study, LD has been considered to be the non-random association of SNP alleles within and between genes. Pairwise LD was computed among the 18 SNPs, and these estimates were compared between the CMS and control Tibetan Chinese group.

The ACE insertion/deletion (rs4340) polymorphism from the RAS system has been extensively studied in HAS. The ACE-I allele has been associated with superior performance in mountaineers who ascend to high altitude compared with the ACE-D allele. It has been found that mountaineers with the ACE-I/I genotype will maintain higher arterial oxygen saturation at rest and during exercise at high altitude than those having the ACE-D/D genotype. Perhaps the genetic reason for these findings is that the 288 bp ACE-I allele generates at least 84 punitive TFBS compared with the ACE-D allele which generates only four TFBS. The implications of this finding have been thoroughly discussed in an earlier paper dealing with acute mountain sickness in Han Chinese. In a previous work, the authors have found that there is a much higher occurrence of ACE-D allele (0.49) in the CMS patients compared with the non-CMS (0.37) Tibetan Chinese control group, again suggesting that the ACE-I allele performs better among individuals in high altitude environments than the ACE-D allele. The angiotensinogen (AGT) M268T polymorphism (rs699) which is also part of the RAS system has been reported to be significantly associated with high-altitude pulmonary edema in Chinese and Indian populations. The AGTR1 polymorphism A1166C (rs5186) which is also part of the RAS system is associated with HAS and other human diseases. These three polymorphisms exhibit changes in LD between their alleles in the CMS study group compared with the non-CMS control group (Figure 1; Table 2), suggesting possible alterations in gene regulation.

Although the APOB polymorphism (rs693) has not been tied to HAS, it has been found to be linked to other human diseases such as dyslipidemia and higher LDL levels and has been shown to influence plasma levels. However, it can be seen that the rs693 APOB SNP exhibits an increase in LD between SNP alleles from the HIF, RAS, and VEGFA pathways (Figure 2; Table 2). The structural protein, apolipoprotein B is a component of LDL and can affect the blood lipid levels by regulating LDL and the balance of cholesterol. It evidently also has some unknown role with the components of HIF, RAS, and VEGFA pathways in CMS cases (Figure 2).

The rs3025039 VEGFA SNP affects the has-miR-591-predicted binding target CA[C/T]GGTC in the 3' UTR, where a VEGFA-C allele would eliminate the binding site for the miRNA. Strong pairwise LD exists between this SNP and the ACE (rs4340), EGNL1 (rs480902), EPAS1, eNOS3 (rs1007311 and rs1799983), and VEGFA (rs699947, rs34357231, rs79469752, and rs2010963) SNPs in the Tibetan control group but disappears in the CMS group (Figure 1). This SNP has also been associated with Behcet's disease, acute myeloid leukemia, Kawasaki disease, psoriatic arthritis, age-related macular degeneration, early-stage non-small-cell lung cancer, and breast cancer.

Future research should involve the ABCGG9 gene, which resides on chromosome 12p and encodes the SUR2 proteins. An RNA splicing variant (SUR2A) of the gene is a cardioprotective ABC protein that serves as a subunit of sarcolemmal
ATP-sensitive K+ channels and is thought to be regulated by hypoxia in high-altitude environments. Intronic SNPs between the 39 exons of the gene have been associated with sleep disorder, depression, hippocampal sclerosis of aging, blood pressure/hypertension, and Hirschsprung disease. It is possible that these SNPs affect BS for the TFs regulating the ABCC9 gene and thereby generate the multiple splicing variants and proteins associated with the gene.

**Conclusion**

Most of the SNPs used in this pairwise LD study have been found to be associated with CMS. These alterations were discussed with regard to changes in gene regulation as a possible cause for CMS. The ACE (I/D) rs4340 SNP would be considered a large contributor to CMS because the ACE-I allele creates many more potential TFBS compared with the ACE-D allele. The ACE-I allele creates a punitive hypoxia-inducible factor 1: aryl hydrocarbon receptor nuclear translocator (HIF1α:ARNT) TFBS that occurs only at this location in the ACE gene. Also, a significant contribution would be the rs3025039 VEGFA SNP whose C allele eliminates the miRNA binding site for miR-591 in the 3'UTR of the gene. The interaction between SNPs within TFBS from different areas of the genome (Table 1) as indicated by LD analysis (Figure 1; Table 2) suggests that allele changes within these regulatory regions may contribute to CMS (Table 2).

**Acknowledgments**

The authors would like to thank Casey Guiland for the technical assistance on the manuscript. It was presented at the 66th Annual Meeting of the American Society of Human Genetics, October 17–22, 2016, Vancouver, Canada. This manuscript is dedicated to the work and memory of Xue-Han Ning.

**Disclosure**

The authors report no conflicts of interest in this work.

**References**


