RESEARCH ARTICLE

Genome-wide analysis of methylation in giant pandas with cataract by methylation-dependent restriction-site associated DNA sequencing (MethylRAD)

Yuyan You1,*, Chao Bai1*, Xuefeng Liu1, Maohua Xia2, Ting Jia1, Xiaoguang Li2, Chenglin Zhang1, Yucun Chen3, Sufen Zhao3, Liqin Wang4, Wei Wang1, Yanqiang Yin5, Yunfang Xiu3, Lili Niu4, Jun Zhou5, Tao Ma2, Yang Du2, Yanhui Liu2

1 Beijing Key Laboratory of Captive Wildlife Technologies, Beijing Zoo, Beijing, China, 2 Beijing Zoo, Beijing, China, 3 Strait (Fuzhou) Giant Panda Research and Exchange Centers, Fuzhou, China, 4 Chengdu Zoo, Chengdu, China, 5 Chongqing Zoo, Chongqing, China

☯ These authors contributed equally to this work.
☆ youyy351@163.com

Abstract

The giant panda (Ailuropoda melanoleuca) is a native species to China. They are rare and endangered and are regarded as the 'national treasure' and 'living fossil' in China. For the time being, there are only about 2500 giant pandas in the world. Therefore, we still have to do much more efforts to protect the giant pandas. In captive wildlife, the cataract incidence of mammalian always increases with age. Currently, in China, the proportion of elderly giant pandas who suffering from cataract has reached 20%. The eye disorder thus has a strong influence on the physical health and life quality of the elderly giant pandas. To discover the genes associated with the pathogenesis of cataract in the elderly giant panda and achieve the goal of early assessment and diagnosis of cataract in giant pandas during aging, we performed whole genome methylation sequencing in 3 giant pandas with cataract and 3 healthy giant pandas using methylation-dependent restriction-site associated DNA sequencing (MethylRAD). In the present study, we obtained 3.62M reads, on average, for each sample, and identified 116 and 242 differentially methylated genes (DMGs) between the two groups under the context of CCGG and CCWGG on genome, respectively. Further KEGG and GO enrichment analyses determined a total of 110 DMGs that are involved in the biological functions associated with pathogenesis of cataract. Among them, 6 DMGs including EEA1, GARS, SLITRK4, GSTM3, CASP3, and EGLN3 have been linked with cataract in old age.

Introduction

Nowadays, a growing number of wild animals have been successfully placed in Zoo. Although the captive animals in Zoo live longer than those in the wild, the aged captive animals (e.g., Malayan Tapir) always encounter various age-related diseases including cataract.
characterized by the opacification of eye lens, is the most common cause for the blindness of almost all mammals, such as dogs, rhesus monkeys, and humans[1–3]. In addition, increasing age is considered to be the most important risk factor for cataract and a considerable number of cataract are classified as age-related cataract[4]. The loss of vision caused by age-related cataract has great influences on the health status of aged animals. As showed by one previous investigation on the captive rhesus monkeys, cataract attacked 20% of the rhesus monkeys at age of 20–22 years and the rate was still increasing after 26 years of age [1]. In addition, the giant panda (Ailuropoda melanoleuca), a world’s most protected rare animals, is also attacked by cataract with age. The studies have shown that the average life span of wild giant pandas is about 15–20 years old, while those in captivity usually live longer and can reach to the age of about 25–30 years[5–6]. Generally, the lifespan of human is 4–4.5-fold longer than the giant panda. The giant panda at the age of 20 years approximately equals human at age of 80–90 years and those aged after 18 years are always served to be aged giant pandas. According to the national survey of eye diseases in aged giant pandas, conducted by Beijing Zoo in 2013, approximately 20% of the aged giant pandas suffered from cataract. Since the giant panda is still an endangered species, the protection of aged giant pandas from cataract has great significances.

Hitherto, a growing number of evidence has shown that genetic factors have large influences on the severity of cataract and play important roles in the development of cataract[7]. For instance, oxidative stress and DNA damage are two common contributors to the many changes in development of age-related cataract[8–10]. Abundant evidence has revealed that genes related to these activities (i.e., oxidative stress and DNA damage) play an important role in the pathogenesis of age-related cataract, such as SOD1, PRDX6, and CRYBA4[11–13]. Epigenetic modifications (e.g., DNA methylation, histone modifications, and non-coding RNA) refer to the alteration of gene activity without any changes in genomic sequence[14–15]. Currently, alteration in epigenetic patterns, in especial DNA methylation, has been closely linked with the cataractogenesis[16]. For example, a reduction of OGG1 and CRYAA expression caused by hypermethylation was observed in lens of eyes with age-related cataract[17–18]. All the existing evidence indicates that the abnormal DNA methylation changes have great contributions to the development of age-related cataract in giant pandas.

In this present study, we performed genome-wide DNA methylation analysis on 3 giant pandas with cataract and 3 healthy giant pandas by methylation-dependent restriction-site associated DNA sequencing (MethylRAD)[19]. Comparison of methylation patterns between the two groups led to the identification of a number of the differentially methylated genes (DMGs) according to the methylation level of CCGG/CCWGG sites. Further analyses showed that the DMGs are preferential located on KEGG pathways and GO terms that have close associations with cataract development. Among these DMGs, some genes (e.g., CASP3, HMGB1, EEA1, and GARS) indeed have been proved to be functioning in the pathogenesis of age-related cataract. Taken together, our study illustrates the epigenetic basis of cataract development in giant panda and identifies potential targets for drug intervention in the therapy of age-related cataract. This research work will facilitate the development of precision medical measures for cataract specific to giant pandas.

**Materials and methods**

**Sampling and MethylRAD sequencing**

The peripheral blood samples were collected from 6 female giant pandas, consisting of 3 cases with cataract and 3 healthy controls. A total of 2 ml blood was draw for each sample during the daily physical examination (without anesthetic). The genomic DNA of blood samples was
extracted using phenol-chloroform method (EMD Millipore-516726, Sigma-Aldrich). Blood samples were initially stored at -80°C. Construction of the MethylRAD library has been described by Wang et al.[19]. 3 μg genomic DNA for each sample was mixed with FspEI (SU/μl) by a volume ratio of 1:0.8. Then, 30 × Enzyme activator was added to the mixture for digestion reaction, with a volume ratio on 0.5:1. After ligation of adaptor, the product was enriched and purified, and then amplified with PCR reactions. PCR product was further purified using QIAquick PCR Purification Kit. Finally, the short DNA fragments in each library were sequenced on Illumina HiSeq platform by the mode of single-end, 50-bp (Illumina Inc., USA).

**Quality control and reads alignment**

To get the clean reads with high quality, we filtered out the poor-quality reads using the threshold of over 15% of bases in a read with quality value of less than 30. In addition, we also removed those reads with a percentage of N greater than 8%[18]. Then, the reads with enzyme sites (enzyme reads) were extracted for subsequent analyses.

The reference genome (AilMel 1.0) of giant panda was downloaded from the National Center for Biotechnology Information (NCBI) with the website: ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCF/000/004/335/GCF_000004335.2_AilMel_1.0/GCF_000004335.2_AilMel_1.0_genomic.fna.gz). Then, we mapped the enzyme reads to the reference genome of AilMel 1.0 using SOAP version 2.21 (http://soap.genomics.org.cn/) with the parameters: -M 4 -v2 -r 0 [20].

**Quantitation and compare of methylation level between two groups**

Since the consistency of amplification efficiency for the sequences with equal length, the methylation level of the sites (CCGG/CCWGG) can be quantified by the sequencing depth of the methylation tag. For the MethylRAD-sequencing, the methylation level of each site (CCGG/CCWGG) was represented by RPM (reads per million) as the following formula[21]:

\[
RPM = \frac{\text{site coverage reads number} \times 1,000,000}{\text{library high quality reads number}}
\]

In addition, the methylation level of one certain genic region including upstream/downstream 2000 bp of TSS (transcription start site), gene body, and upstream/downstream 2000 bp of TTS (transcription termination site) was calculated by the sum value of all the methylated sites that are located in the corresponding the region. The methylation data were then analyzed to identify the differentially methylated sites/genes (DMS/Gs) between the case and control groups using the edgeR Bioconductor package that is relied on the number of coverage reads on the sites or genes[22]. Here, the sites/genes with at least 3 reads coverage across the samples in at least one group were retained and those meeting the threshold of p value ≤ 0.05 and \(|\log2FC| > 1\) were determined as DMS/Gs, with hypermethylated and hypomethylated.

**Gene annotation and enrichment analysis**

The gene annotation information of giant panda was also downloaded from the National Center for Biotechnology Information (NCBI) with the website: ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCF/000/004/335/GCF_000004335.2_AilMel_1.0/GCF_000004335.2_AilMel_1.0_genomic.gff.gz. The UTR regions of the genes were calculated by using SnpEff tool based on the annotation information[23]. The distributions of the methylated sites on various genomic sequences were calculated by BEDTools[24]. To perform gene set enrichment analysis, we obtained the available gene information of pathways and biological functions from databases.
of Kyoto Encyclopedia of Genes and Genomes (KEGG), Gene Ontology (GO) and The Comparative Toxicogenomics Database (CDT)[25–27]. We utilized the hypergeometric test to calculate the statistical significance of genes enriched on each biological function.

Results

Source of the study samples

In 2013, we looked into 55 old giant pandas in Chinese Zoo, among which 11 (8 females and 3 males) were suffering from cataract. Most of the sufferers were female and over 20 years old. Here, we obtained the genomic DNA samples of peripheral blood cells from 3 female giant pandas with cataract, 2 healthy female giant pandas and 1 male giant panda to perform subsequent genome-wide methylation study. As presented in Table 1, we numbered the 6 giant panda samples as YY-Y, BD-Y, YY-XK, LL-D, JN-D, and XX-XK, respectively. Among these samples, the YY-Y and BD-Y, were healthy ones with an age of approximately 20 years old, YY-XK was healthy with an age of 29 years old. LL-D, JN-D, and XX-XK were ill ones with an age of 36, 32, and 25 years old, respectively. LL-D was died in 2018.

MethylRAD of the samples

The methylomes of the peripheral blood cells from the 6 giant pandas were generated on HiSeq platform using MethylRAD sequencing (see Materials and Methods)[19]. Here, we obtained 3.62 ± 0.17 million raw sequencing reads, on average, for each sample. After removing the reads with low quality and the reads without enzyme sites, we retained about 1.72 million clean enzyme (FspEi) reads that covering CCGG/CCWGG sites in each sample for the subsequent analyses. We mapped the enzyme reads to the reference genome (AilMel 1.0) of giant panda by using SOAP software version 2.21[20]. On average, about 1.31 million clean reads were mapped to the genome for each sample, the mapping ratio is about 76.66% (Table 2). In this study, we determined the reliable methylated sites by a cutoff of read-coverage no less than 3, and obtained 1 million CCGG sites and 0.32 million CCWGG sites, on average, for each sample. The average coverage depth of CCGG and CCWGG sites is 10.3 and 9.01, respectively (Table 3). Therefore, the sequencing reads satisfy the condition of following analyses.

Signatures of DNA methylation in the giant pandas

Then, we analyzed the distribution of the methylated CCGG/CCWGG sites on distinct genomic sequences, including Utr3prime, Utr5prime, Upstream, Exon, Intron, and intergenic regions (Fig 1). Among them, the sites were most located in intergenic and intron regions, the next in exon and upstream, and the least in Utr3prime and Utr5prime regions. In addition, results revealed that the distribution of methylated CCGG and CCWGG sites on various

Table 1. Basic characteristics of giant pandas.

| Name   | Spectrum number | Number       | Birth year | Status        | Sex     | Remarks       |
|--------|-----------------|--------------|------------|---------------|---------|---------------|
| YAER   | 493             | YY-X         | 1999       | Health        | Female  | -             |
| BINGDIAN | 520          | BD-X         | 2000       | Health        | Male    | -             |
| YAYA   | 362             | YY-XK        | 1990       | Health        | Female  | -             |
| LELE   | 320             | LL-D         | 1986       | Age-related cataract | Female | DEATH         |
| JINI   | 403             | JN-D         | 1993       | Age-related cataract | Female | -             |
| XINXING | 253            | XX-XK        | 1982       | Age-related cataract | Female | -             |

https://doi.org/10.1371/journal.pone.0222292.t001
Genomic regions was very similar in the 3 healthy giant pandas, while a marked difference was observed in the 3 giant panda that suffering from cataract. For instance, the number of methylated CCGG and CCWGG sites located in the various genomic regions was fewest in LL-D, moderate in XX-XK, and largest in JN-D. In addition, we found that the methylated CCGG and CCWGG sites in LL-D and XX-XK were smaller than the healthy giant pandas. However, for the methylated sites in JN-D, the number of CCGG sites was larger than healthy giant pandas, while the number of CCWGG sites was similar with healthy giant pandas.

In addition, we analyzed the overall methylation pattern of different positions on the genic regions. As presented in Fig 2 and S1 Table, the methylation level based on CCGG and CCWGG sites was gradually rising from the initial position of genic region, occurred a turning point of rising tread, then continued to rise and reach the highest value at the end position. Moreover, no obvious differences were observed in the 6 giant pandas.

Genes differentially methylated in the case and control groups

In the present study, we calculated the methylation level of genes based on the CCGG and CCWGG sites, respectively, and identified the DMGs between the case and control groups using a threshold of \( P \) value \( \leq 0.05 \) and absolute log2FC value > 1 (see Materials and Methods). Here, we identified a total of 116 DMGs by the CCGG sites, including 75 hypermethylated genes and 41 hypomethylated genes. Moreover, we determined 242 DMGs by the CCWGG sites, including 164 hypermethylated genes and 78 hypomethylated genes. The heat-map plot showed a different methylation pattern between the two groups, with genes highly methylated in patients presenting low methylation level in the healthy samples and vice versa. For both the CCWGG and CCGG sites, there were a large number of genes with hypermethylation in healthy group (Fig 3). Among them, there were 20 DMGs that were identified by both the CCGG and CCWGG sites, containing 12 hypermethylated genes and 8 hypomethylated genes (Fig 4 and S2 Table).
Pathway-level functions of the DMGs in the development of cataract

To determine the signaling pathways potentially associated with age-related cataract, we first performed KEGG enrichment analysis and searched the available annotation information from CTD database[25, 27]. By using a threshold of \( P \) value \(<0.05\), we identified 15 and 55 enriched KEGG signaling pathways for the CCGG- and CCWGG-based DMGs, respectively. Among them, the CCGG-based signaling pathways contained 27 DMGs, including 22 hypermethylated genes and 5 hypomethylated genes; while the CCWGG-based signaling pathways contained 96 DMGs, including 80 hypermethylated genes and 16 hypomethylated genes. In addition, there were 3 enriched pathways according to both CCGG- and CCWGG-based DMGs. A total of 10 DMGs were located in these 3 signaling pathways, including 8 hypermethylated genes and 2 hypomethylated genes. We then summarized the findings of KEGG signaling pathways to assess their associations with cataract pathogenesis. The pathways associated with genetic information processing included base excision repair (cataract-related genes: \( HMGB1 \), hypomethylated in aged giant pandas with cataract), SNARE interactions in vesicular transport (\( STX19 \), hypermethylated), and RNA degradation (\( MPHOSPH6 \) and \( TTC37 \), both were hypermethylated). The pathways related with environmental information processing contained NF-kappa B signaling pathway (\( CCL19 \), hypomethylated), cAMP signaling pathway, HIF-1 signaling pathway (\( LOC100484901 \), \( PDK1 \), and \( EGLN3 \), all were hypermethylated). On the level of cellular process, there were 5 pathways, 3 of which had some associations with cataract, including cell cycle (\( ORC6 \) and \( CDC7 \), both were hypermethylated), apoptosis (\( CASP3 \), hypermethylated), p53 signaling pathway (\( TP53I3 \) and \( CASP3 \), both were hypermethylated). On the level of metabolism, there were 17 pathways, 7 of which were related with cataract, including drug metabolism-cytochrome P450 (\( FMO5 \) and \( GSTM3 \), both were hypermethylated), glycerolipid metabolism (\( LPL \), hypermethylated), beta-Alanine metabolism (\( LOC100474209 \), hypermethylated), tyrosine metabolism (\( LOC100474209 \), hypermethylated),
phenylalanine metabolism (LOC100474209, hypermethylated), metabolism of xenobiotics by cytochrome P450 (GSTM3, hypermethylated), steroid biosynthesis (MSMO1, hypermethylated). For organismal systems, there were 16 pathways, 1 of which was related with cataract (i.e., axon guidance). For the level of human disease, there were 19 pathways, 5 of which were associated with cataract. Those were Epithelial cell signaling in Helicobacter pylori infection (ATP6V1C1 and CASP3, both were hypermethylated), Platinum drug resistance (TOP2B, GSTM3, and CASP3, all were hypermethylated), Viral myocarditis (LOC100463889 and CASP3, both were hypermethylated), Fluid shear stress and atherosclerosis (LOC100484313, GSTM3, and ACVR2A, all were hypermethylated), Chemical carcinogenesis (GSTM3, hypermethylated) (Fig 5, Table 4 and S3 Table).

GO-level functions of the DMGs in cataract

Similarly, we also conducted GO enrichment analysis for the DMGs and then found the cataract-related GO term by CTD database[26–27]. The CCGG-based DMGs were enriched in 396 GO terms, containing 59 DMGs (36 hypermethylated genes, 23 hypomethylated genes). The CCWGG-based DMGs were enriched in 681 GO terms, including 129 DMGs (88 hypermethylated genes and 41 hypomethylated genes). There were 105 GO terms that were identified with both CCGG- and CCWGG-based DMGs. These terms contained 59 DMGs, among which 7 genes carried both differentially methylated CCGG and CCWGG sites. Under the cellular component, there were 102 enriched terms, among which 26 terms had an association with cataract. In addition, there were 13 terms containing DMGs with same directional
**Fig 3.** Cluster heat map of differentially methylated genes between group.

https://doi.org/10.1371/journal.pone.0222292.g003

**Fig 4.** Gene statistics of different methylation levels.

https://doi.org/10.1371/journal.pone.0222292.g004
methylation changes, such as intermediate filament (LOC105240942 and LOC100465932, hypermethylated), membrane raft (RRK2, LOC100476759, and GNAI1, hypermethylated), midbody (KIF20B, GNAI1, and ASPM, hypermethylated). For molecular function, the DMGs were enriched in 181 terms, among which 29 were associated with cataract. 16 of the 29 terms had DMGs showing same directional changes, such as structural constituent of cytoskeleton (LOC100473181 and TUBD1, hypermethylated), SNARE binding (STX19 and LRRK2, hypermethylated), beta-catenin binding (SOX9 and SOX17, hypomethylated), calmodulin binding (EEA1 and MIP, hypermethylated). For biological process, the DMGs were enriched in 690 terms, 219 of which were linked with cataract. 182 of the 219 terms had DMGs showing same directional changes, such as cardiac muscle tissue development (NKX2-5, hypermethylated), BMP signaling pathway (ACVR2A, TMEM100, and NKX2-5, hypermethylated), cellular metabolic process (PDP1 and PDK1, hypermethylated), cholesterol metabolic

Fig 5. Functional enrichment of genes related to different methylation levels.

https://doi.org/10.1371/journal.pone.0222292.g005
process (PCTP and LOC100476613, hypomethylated), and regulation of membrane potential (GABRG1, LRRK2, and LOC105234775, hypermethylated) (Fig 5, Table 5 and S4 Table).

Roles of the DMGs in cataract

In this study, we identified a total of 338 DMGs between the groups. Among the DMGs, 116 have been previously supposed to be potentially linked with the development of cataract (S5 Table). Based on the results of enrichment analysis, we selected 16 DMGs associated with the cataract-related KEGG pathways and 108 DMGs involved in the cataract-related GO terms. By combing the results of KEGG and GO enrichment analyses, we obtained a total of 110 DMGs (Table 6), among which 6 have been linked with the cataract in old age in previous reports that were EEA1, GARS, SLITRK4, GSTM3, CASP3, and EGLN3.

Discussion

In the present study, we analyzed the methylation profile differences between the aged giant pandas suffering from cataract and healthy giant pandas and identified hundreds of DMGs in giant pandas with age-related cataract. Notably, we found no methylation differences on the genes (i.e., GLB1, CDKN2A and CDKN2B) highly correlated with aging[28–29], implying negligible effects of age on the DMGs. Further analysis showed that the genes with significant methylation differences between the case and control groups are indeed located on many biological processes related with cataract formation, such as base excision repair, p53 signaling pathway, and apoptosis[18, 30, 31]. Among them, p53 signaling pathway plays an important roles in the prevention of apoptosis of lens epithelial cells and cataractogenesis[31], the hypermethylation of genes (i.e., TP53I3 and CASP3) on this pathway would like to downregulate the
| CClass | GO_id   | GO_def                   | CCGG_pval | CCWGG_pval |
|--------|---------|--------------------------|-----------|------------|
| Cellular_component | GO:0005882 | intermediate filament | 0.026127779 | 0.02394968 |
|        | GO:0005667 | transcription factor complex | 0.016392771 | 0.010378895 |
|        | GO:0016607 | integral component of membrane | 0.009779408 | 0.90683214 |
|        | GO:0016623 | nuclear speck | 0.032533867 | 0.24309679 |
|        | GO:0005645 | cell | 0.018123614 | NA |
|        | GO:0005615 | extracellular space | 0.03795066 | 0.376522069 |
|        | GO:0003018 | Z-disc | 0.049478525 | 0.218824299 |
|        | GO:0015102 | extracellular matrix | 0.038135622 | 0.14389526 |
|        | GO:0016636 | nuclear matrix | 0.038135622 | NA |
|        | GO:001903 | early endosome membrane | 0.0046239 | 0.20585267 |
|        | GO:0028399 | dendrite cytoplasm | NA | 0.00426355 |
|        | GO:0043204 | perikaryon | NA | 0.00424455 |
|        | GO:0034266 | growth cone | NA | 0.0446993 |
|        | GO:0000474 | chromosome, telomeric region | NA | 0.02660753 |
|        | GO:0049005 | neuron protection | NA | 0.040183436 |
|        | GO:0045123 | membrane raf | NA | 0.02576856 |
|        | GO:0030666 | endocytic vesicle membrane | NA | 0.029466371 |
|        | GO:0044291 | cell-cell contact zone | NA | 0.00426365 |
|        | GO:0034946 | mitochondrion | NA | 0.025032919 |
|        | GO:0005730 | nucleolus | 0.26912545 | 0.029526994 |
|        | GO:0005759 | mitochondrial matrix | NA | 3.68158E-05 |
|        | GO:0005921 | gap junction | NA | 0.08306393 |
|        | GO:0005637 | nuclear inner membrane | NA | 0.038072518 |
|        | GO:0042645 | mitochondrial nucleoid | NA | 0.04490081 |
|        | GO:0031233 | spindle midzone | NA | 0.01801158 |
|        | GO:0031410 | cytoplasmic vesicle | NA | 0.02437127 |
| Molecular_function | GO:0043665 | sequence-specific DNA binding | 0.010490985 | 0.04697174 |
|        | GO:0044998 | protein dimerization activity | 0.00322677 | 0.01831505 |
|        | GO:0001228 | transcriptional activator activity, RNA polymerase II transcription regulatory region sequence-specific binding | 0.001586706 | 0.019205381 |
|        | GO:0003690 | double-stranded DNA binding | 0.0038267 | NA |
|        | GO:0005200 | structural constituent of cytoskeleton | 0.028518085 | 0.13707187 |
|        | GO:0004149 | SNAPE binding | 0.00981659 | 0.05239303 |
|        | GO:0007000 | transcription factor activity, sequence-specific DNA binding | 0.043134111 | 0.610799417 |
|        | GO:000735 | structural constituent of ribosome | 0.043990434 | 0.305199845 |
|        | GO:0046992 | protein heterodimerization activity | 0.015107415 | 0.13882043 |
|        | GO:0004129 | cytochrome-c oxidase activity | 0.003867281 | NA |
|        | GO:0001077 | transcriptional activator activity, RNA polymerase II core promoter proximal region sequence-specific binding | 0.0021344047 | 0.43540707 |
|        | GO:0003682 | chromatin binding | 0.0283346 | 0.547748796 |
|        | GO:0008013 | beta-catenin binding | 0.014292521 | 0.074006113 |
|        | GO:0002576 | protein disulfide isomerase activity | NA | 0.010811581 |
|        | GO:0016740 | transerase activity | NA | 0.02660753 |
|        | GO:0005056 | iron ion binding | NA | 0.02576856 |
|        | GO:0034148 | L-ascorbic acid binding | NA | 0.015108679 |
|        | GO:0016301 | kinase activity | NA | 0.002936508 |
|        | GO:0003432 | ferrooxidase activity | NA | 0.003395245 |
|        | GO:0005516 | calmodulin binding | NA | 0.00899542 |
|        | GO:0005215 | transporter activity | NA | 0.012989815 |
|        | GO:000254 | C-4 methylsterol oxidase activity | NA | 0 |
|        | GO:0031059 | NF-kappaB binding | NA | 0.00306393 |
|        | GO:0004364 | glutathione transferase activity | NA | 0.020016375 |
|        | GO:0004820 | CTD phosphatase activity | NA | 0.02660753 |
|        | GO:0004672 | protein kinase activity | 0.080285293 | 0.218824299 |
|        | GO:0015250 | water channel activity | NA | 0.003395245 |
|        | GO:0002122 | structural constituent of eye lens | NA | 0.01831505 |
|        | GO:0002166 | nucleotide binding | NA | 0.019035545 |

(Continued)
| CClass          | GO_id     | GO_def                                      | CCGG_pval   | CCWGG_pval   |
|-----------------|-----------|---------------------------------------------|-------------|--------------|
| Biological.process | 0040738   | cardiac muscle tissue development           | 0.000445093 | 0.00265811   |
|                 | 0055007   | cardiac muscle cell differentiation         | 0.004207877 | 0.02360783   |
|                 | 00007283  | spermatogenesis                             | 0.0051856   | 0.04069737   |
|                 | 0001570   | vasculogenesis                              | 0.0103349   | 0.00059643   |
|                 | 00360821  | protein stabilization                       | 0.040750902 | 0.04436998   |
|                 | 0090090   | negative regulation of canonical Wnt signaling pathway | 0.004784129 | 0.000191203 |
|                 | 0060047   | heart contraction                           | 0.00867457  | 0.00519712   |
|                 | 0060038   | cardiac muscle cell proliferation           | 0.00239694  | 0.00442798   |
|                 | 0030509   | BMP signaling pathway                       | 0.012527107 | 0.000792502  |
|                 | 0033007   | heart morphogenesis                         | 0.003867281 | 0.021781089  |
|                 | 00003161  | cardiac conduction system development       | 9.63811E-05 | 0.000584755  |
|                 | 0031295   | T cell costimulation                        | 0.003867281 | 0.00486876   |
|                 | 00355008  | cardiac muscle tissue morphogenesis         | 0.000334894 | 0.002006702  |
|                 | 0043491   | protein kinase B signaling                  | 0.002609396 | 0.013606467  |
|                 | 0043860   | positive regulation of protein kinase activity | 0.006961412 | 0.038072128  |
|                 | 0070328   | triglyceride homeostasis                    | 0.001638839 | 0.009521899  |
|                 | 00355050  | embryonic heart tube development            | 0.00867457  | 0.007151247  |
|                 | 1901203   | positive regulation of extracellular matrix assembly | 4.83169E-05 | 0.000294308  |
|                 | 0060048   | cardiac muscle contraction                  | 0.005309692 | 0.029446371  |
|                 | 00003221  | right ventricular cardiac muscle tissue morphogenesis | 1.61479E-05 | 9.87507E-05  |
|                 | 0034504   | protein localization to nucleus             | 0.00322772  | 0.01835105   |
|                 | 0010989   | positive regulation of sequestering of triglyceride | 0.00239694  | 0.00442798   |
|                 | 1903779   | regulation of cardiac conduction            | 1.61479E-05 | 9.87507E-05  |
|                 | 0007017   | microtubule-based process                   | 0.005703272 | 0.011507479  |
|                 | 0046330   | positive regulation of JNK cascade          | 0.008334034 | 0.049900881  |
|                 | 0008284   | positive regulation of cell proliferation   | 0.014966279 | 0.402952797  |
|                 | 0030857   | negative regulation of epithelial cell differentiation | 0.000160215 | NA           |
|                 | 0014068   | positive regulation of phosphatidylinositol 3-kinase signaling | 0.000105939 | NA           |
|                 | 00069354  | inflammatory response                       | 0.001597038 | 0.158902846  |
|                 | 0043065   | positive regulation of apoptotic process    | 0.025787421 | NA           |
|                 | 0030154   | cell differentiation                        | 0.00966065  | 0.193072503  |
|                 | 00001934  | positive regulation of protein phosphorylation | 0.033549391 | 0.157750125  |
|                 | 0007507   | heart development                           | 0.006927629 | 0.071080057  |
|                 | 0009408   | response to heat                            | 0.003867281 | NA           |
|                 | 0010458   | exit from mitosis                           | 0.003343694 | NA           |
|                 | 0010628   | positive regulation of gene expression      | 0.012535059 | 0.113861252  |
|                 | 0014032   | neural crest cell development               | 0.000334694 | NA           |
|                 | 0030903   | notochord development                       | 0.000160215 | NA           |
|                 | 0070830   | bicellular tight junction assembly          | 0.00652924  | NA           |
|                 | 0031897   | positive regulation of protein kinase B signaling | 0.017457942 | NA           |
|                 | 0060009   | Sertoli cell development                    | 0.000334694 | NA           |
|                 | 0070374   | positive regulation of ERK1 and ERK2 cascade | 0.038996766 | NA           |
|                 | 0031532   | actin cytoskeleton reorganization            | 0.01862701  | NA           |
|                 | 0010942   | positive regulation of cell death           | 0.001038232 | NA           |
|                 | 0032868   | response to insulin                         | 0.004207877 | NA           |
|                 | 0001502   | cartilage condensation                      | 0.001038232 | NA           |
|                 | 0032757   | positive regulation of interleukin-8 production | 0.001233799 | NA           |
|                 | 0090023   | positive regulation of neutrophil chemotaxis | 0.000867457 | NA           |
|                 | 0008584   | male gonad development                      | 0.011407572 | 0.060195631  |
|                 | 0050892   | cell morphogenesis                          | 0.01862701  | NA           |
|                 | 0037090   | regulation of catalytic activity            | 0.012527107 | NA           |
|                 | 0007417   | central nervous system development          | 0.034419441 | NA           |
|                 | 0045666   | positive regulation of neuron differentiation | 0.020184093 | 0.101044793  |
|                 | 0071560   | cellular response to transforming growth factor beta stimulus | 0.004561885 | NA           |
|                 | 0001494   | tissue homeostasis                          | 0.00607067  | NA           |
|                 | 0032735   | positive regulation of interleukin-12 production | 0.000867457 | NA           |

(Continued)
Table 5. (Continued)

| CClass   | GO_id       | GO_def                                      | CCGG_pval  | CCWGG_pval |
|----------|-------------|---------------------------------------------|------------|------------|
| GO:0030679 | positive regulation of epithelial cell proliferation | 0.006961412 | NA         |
| GO:0030858 | positive regulation of epithelial cell differentiation | 9.63811E-05 | NA         |
| GO:0046931 | positive regulation of mitotic cell cycle | 0.002111633 | NA         |
| GO:0032496 | response to lipopolysaccharide | 0.026127779 | NA         |
| GO:0014911 | positive regulation of smooth muscle cell migration | 0.00168839 | NA         |
| GO:0006935 | immune response | 0.034996371 | 0.259172857 |
| GO:0070168 | negative regulation of biomineral tissue development | 1.61479E-05 | NA         |
| GO:0071260 | cellular response to mechanical stimulus | 0.007406235 | NA         |
| GO:0007257 | activation of IJN kinase activity | 0.004929193 | NA         |
| GO:0010629 | negative regulation of gene expression | 0.04649987 | NA         |
| GO:0030097 | hemopoiesis | 0.012527107 | 0.065604136 |
| GO:0030727 | regulation of inflammatory response | 0.006109824 | NA         |
| GO:0030335 | positive regulation of cell migration | 0.049478525 | NA         |
| GO:0032760 | positive regulation of tumor necrosis factor production | 0.00013679 | NA         |
| GO:0043893 | positive regulation of transcription, DNA-templated | 0.048018346 | 0.24856192 |
| GO:0071564 | cellular response to epidermal growth factor stimulus | 0.001424041 | NA         |
| GO:0030879 | mammary gland development | 0.001038232 | NA         |
| GO:0060174 | limb bud formation | 0.000445093 | NA         |
| GO:002062 | chondrocyte differentiation | 0.004207877 | NA         |
| GO:0060174 | skeletal system development | 0.039821147 | NA         |
| GO:0032755 | positive regulation of interleukin-6 production | 0.00350208 | NA         |
| GO:0043123 | positive regulation of PLC-pka/A kinase/NF-kappaB signaling | 0.006502616 | 0.062602373 |
| GO:0032732 | positive regulation of interleukin-1 production | 0 | NA         |
| GO:006935 | chemotaxis | 0.029333711 | 0.14047622 |
| GO:0030776 | regulation of immune response | 0.007406235 | NA         |
| GO:0030718 | positive regulation of interleukin-1 beta secretion | 0.006711596 | NA         |
| GO:0020020 | positive regulation of male gonad development | 9.63811E-05 | NA         |
| GO:0032332 | positive regulation of chondrocyte differentiation | 0.006711596 | NA         |
| GO:0007186 | G-protein coupled receptor signaling pathway | 0.001350659 | 0.352837434 |
| GO:0031216 | cartilage development | 0.00902699 | NA         |
| GO:0018357 | epithelial to mesenchymal transition | 0.00168839 | NA         |
| GO:006913 | apoptotic process | 0.018195573 | 0.08355214 |
| GO:0048469 | cell maturation | 0.004207877 | NA         |
| GO:0007517 | muscle organ development | 0.019486784 | NA         |
| GO:007626 | locomotory behavior | 0.023824179 | NA         |
| GO:0001503 | ossification | 0.020184093 | NA         |
| GO:0003388 | chromatin remodeling | 0.002314778 | NA         |
| GO:0006309 | apoptotic DNA fragmentation | 0.006502616 | NA         |
| GO:0001541 | ovarian follicle development | 0.00322677 | NA         |
| GO:0007010 | cytoskeleton organization | 0.042634951 | NA         |
| GO:0016042 | lipid catabolic process | 0.031835591 | 0.150790378 |
| GO:0010976 | positive regulation of neuron projection development | 0.008126164 | NA         |
| GO:0071300 | cellular response to retinoic acid | 0.01961532 | NA         |
| GO:0007595 | lactation | 0.002111633 | NA         |
| GO:0045807 | positive regulation of endocytosis | 0.000867457 | NA         |
| GO:0071599 | ovot vesicle development | 0.006502616 | NA         |
| GO:0031175 | neuron projection development | 0.039823147 | 0.182544352 |
| GO:0047372 | positive regulation of protein catabolic process | 0.014292521 | 0.070406113 |
| GO:0032436 | positive regulation of proteasomal ubiquitin-dependent protein catabolic process | NA | 0.00630072 |
| GO:0092001 | negative regulation of release of cytochrome c from mitochondria | NA | 0.00421635 |
| GO:0044237 | cellular metabolic process | NA | 0.0012356 |
| GO:0014044 | Schwann cell development | NA | 0.00384755 |
| GO:0031092 | positive regulation of NF-kappaB transcription factor activity | NA | 0.00670258 |
| GO:0016310 | phosphorylation | NA | 0.00168676 |
| GO:0030316 | osteoclast differentiation | NA | 0.013866467 |
| GO:0030848 | regulation of calcium-mediated signaling | NA | 0.00986205 |

(Continued)
| CClass | GO_id | GO_def | CCGG_pval | CCWGG_pval |
|--------|-------|--------|-----------|------------|
| GO:0042789 | mRNA transcription from RNA polymerase II promoter | NA | 0.003395245 |
| GO:0002376 | immune system process | NA | 0.020163575 |
| GO:0031402 | neuron apoptotic process | NA | 0.001486876 |
| GO:0070509 | calcium ion import | NA | 0.02360783 |
| GO:0070940 | dephosphorylation of RNA polymerase II C-terminal domain | NA | 0.000968205 |
| GO:0030999 | regulation of nitric-oxide synthase activity | NA | 0.004216355 |
| GO:0032967 | positive regulation of collagen biosynthetic process | NA | 0.005119712 |
| GO:006090 | pyruvate metabolic process | NA | 0.000494847 |
| GO:0072593 | reactive oxygen species metabolic process | NA | 0.02360783 |
| GO:0071773 | cellular response to BMP stimulus | NA | 0.015108679 |
| GO:1900117 | positive regulation of cytokine production involved in inflammatory response | NA | 0.003395245 |
| GO:1901800 | positive regulation of proteasomal protein catabolic process | NA | 0.005119712 |
| GO:0070997 | neuron death | NA | 0.005119712 |
| GO:0018015 | peptidyl-serine phosphorylation | NA | 0.029703889 |
| GO:0043525 | positive regulation of neuron apoptotic process | NA | 0.000968205 |
| GO:0043537 | negative regulation of blood vessel endothelial cell migration | NA | 0.002006702 |
| GO:0030239 | myosinfil assembly | NA | 0.004216355 |
| GO:0008203 | cholesterol metabolic process | NA | 0.003453434 |
| GO:0032092 | positive regulation of protein binding | NA | 0.003360189 |
| GO:0014937 | negative regulation of endothelial cell proliferation | NA | 0.00096502 |
| GO:0030889 | negative regulation of B cell proliferation | NA | 0.003395245 |
| GO:006919 | heart formation | NA | 0.004247298 |
| GO:0007013 | actin filament organization | NA | 0.010784844 |
| GO:0021766 | hippocampus development | NA | 0.042616389 |
| GO:006914 | autophagy | NA | 0.040285359 |
| GO:006006 | glucose metabolic process | NA | 0.003797886 |
| GO:0090263 | positive regulation of canonical Wnt signaling pathway | NA | 0.030403443 |
| GO:0001615 | MAPK cascade | 0.079101181 | 0.02803752 |
| GO:0008853 | carnitine shuttle | NA | 0.004216355 |
| GO:0060135 | maternal process involved in female pregnancy | NA | 0.002006702 |
| GO:007080 | mitotic metaphase plate congression | NA | 0.038017218 |
| GO:006633 | fatty acid beta-oxidation | NA | 0.040219697 |
| GO:0010718 | positive regulation of epithelial to mesenchymal transition | NA | 0.012173591 |
| GO:0008340 | determination of adult lifespan | NA | 0.002006702 |
| GO:0048589 | developmental growth | NA | 0.015180679 |
| GO:0070035 | mitochondrion organization | NA | 0.014271521 |
| GO:007040 | lysosome organization | NA | 0.02549204 |
| GO:0016525 | negative regulation of angiogenesis | NA | 0.003453434 |
| GO:0007214 | gamma-aminobutyric acid signaling pathway | NA | 0.00096502 |
| GO:0031389 | negative regulation of protein ubiquitination | NA | 0.02549204 |
| GO:0008219 | cell death | NA | 0.018351505 |
| GO:0007049 | cell cycle | NA | 0.008002539 |
| GO:0030877 | neurological system process | NA | 0.010154547 |
| GO:0048812 | neuron projection morphogenesis | NA | 0.004637323 |
| GO:0021549 | cerebellum development | NA | 0.0294637 |
| GO:0042220 | response to cocaine | NA | 0.005119712 |
| GO:0042752 | regulation of circadian rhythm | NA | 0.042616389 |
| GO:0070519 | skeletal muscle tissue development | NA | 0.00242571 |
| GO:007200 | renal system development | NA | 0.002006702 |
| GO:008016 | regulation of heart contraction | NA | 0.009521989 |
| GO:0030896 | response to stimulus | NA | 0.016353363 |
| GO:1902476 | chloride transmembrane transport | NA | 0.02949468 |
| GO:008833 | water transport | NA | 0.004637323 |
| GO:0042593 | glucose homeostasis | NA | 0.02949468 |
| GO:0032091 | negative regulation of protein binding | NA | 0.040219697 |
| GO:1903215 | negative regulation of protein targeting to mitochondrion | NA | 0.002943408 |

(Continued)
Table 5. (Continued)

| CClass | GO_id     | GO_def                                      | CC{ }GG_pval | CW{ }GG_pval |
|--------|-----------|---------------------------------------------|--------------|--------------|
|        | GO:0001816 | cytokine production                         | NA           | 0.004216355  |
|        | GO:2000484 | positive regulation of interleukin-8 secretion | NA           | 0.000584755  |
|        | GO:0046848 | positive regulation of erythrocyte differentiation | NA           | 0.00806393   |
|        | GO:0009566 | fertilization                               | NA           | 0.03362847    |
|        | GO:0042713 | sperm ejaculation                            | NA           | 9.87507E-05   |
|        | GO:0008631 | intrinsic apoptotic signaling pathway in response to oxidative stress | NA           | 0.0263811    |
|        | GO:0071158 | positive regulation of cell cycle arrest     | NA           | 0.006103617   |
|        | GO:2000573 | positive regulation of DNA biosynthetic process | NA           | 0.001442798   |
|        | GO:0009777 | DNA damage response, signal transduction by p53 class mediator resulting in cell cycle arrest | NA           | 0.027441835   |
|        | GO:006096  | glycolytic process                          | NA           | 0.031507479   |
|        | GO:0021987 | cerebral cortex development                 | NA           | 0.00329129    |
|        | GO:0002444 | hematopoietic progenitor cell differentiation | NA           | 0.018325705   |
|        | GO:2000469 | negative regulation of peroxidase activity  | NA           | 0             |
|        | GO:0016126 | sterol biosynthetic process                 | NA           | 0.0265831    |
|        | GO:0019915 | lipid storage                               | NA           | 0.016678719   |
|        | GO:0032016 | keratinocyte differentiation                 | NA           | 0.02946371    |
|        | GO:0031091 | positive regulation of sequence-specific DNA binding transcription factor activity | NA           | 0.00401609  |
|        | GO:0006633 | fatty acid biosynthetic process             | NA           | 0.047411578   |
|        | GO:0033038 | negative regulation of cell growth          | NA           | 0.03767988    |
|        | GO:0008357 | regulation of transcription from RNA polymerase II promoter | 0.102048797 | 0.021169358  |
|        | GO:0046678 | negative regulation of insulin secretion    | NA           | 0.01835105    |
|        | GO:1902236 | negative regulation of endoplasmic reticulum stress-induced intrinsic apoptotic signaling pathway | NA           | 0.007166394   |
|        | GO:0045665 | negative regulation of neuron differentiation | NA           | 0.00733058    |
|        | GO:0010906 | regulation of glucose metabolic process     | NA           | 0.00806393    |
|        | GO:1901213 | negative regulation of neuron death         | NA           | 0.000151247   |
|        | GO:0031398 | positive regulation of protein ubiquitination | NA           | 0.00428969    |
|        | GO:0071346 | cellular response to interferon-gamma       | NA           | 0.027441835   |
|        | GO:0071850 | mitotic cell cycle arrest                    | NA           | 7.8137E-05    |
|        | GO:0030713 | positive regulation of cytokine secretion   | NA           | 0.00806393    |
|        | GO:0009356 | complement activation                       | NA           | 0.000584755   |
|        | GO:0045840 | positive regulation of mitotic nuclear division | NA           | 0.006103617   |
|        | GO:0048144 | behavioral response to cocaine              | NA           | 0.001442798   |
|        | GO:0010508 | positive regulation of autophagy            | NA           | 0.027441835   |
|        | GO:0032555 | response to estradiol                       | NA           | 0.047411578   |
|        | GO:0045997 | positive regulation of cell differentiation  | NA           | 0.01835105    |
|        | GO:0022038 | corpus callosum development                 | NA           | 0.00098205    |
|        | GO:0031646 | mitochondrion localization                  | NA           | 0.001442798   |
|        | GO:0019722 | calcium-mediated signaling                  | NA           | 0.047411578   |
|        | GO:0031641 | locomotory exploration behavior             | NA           | 0.003395245   |
|        | GO:0090394 | negative regulation of excitatory postsynaptic potential | NA           | 0.002006702   |
|        | GO:0048167 | regulation of synaptic plasticity           | NA           | 0.02036375    |
|        | GO:0030182 | neuron differentiation                      | NA           | 0.044306918   |
|        | GO:2001214 | positive regulation of vasculogenesis       | NA           | 0.00265811    |
|        | GO:0043409 | negative regulation of MAPK cascade         | NA           | 0.001442798   |
|        | GO:1900701 | positive regulation of ER-associated ubiquitin-dependent protein catabolic process | NA           | 0.003395245   |
|        | GO:1901741 | positive regulation of myoblast fusion      | NA           | 0.00519712    |
|        | GO:0006749 | glutathione metabolic process               | NA           | 0.02946371    |
|        | GO:0048514 | blood vessel morphogenesis                  | NA           | 0.006103617   |
|        | GO:0042391 | regulation of membrane potential            | NA           | 0.012989815   |
|        | GO:0046855 | inositol phosphate dephosphorylation        | NA           | 0.00265811    |
|        | GO:1901214 | regulation of neuron death                  | NA           | 0.006103617   |

https://doi.org/10.1371/journal.pone.0222292.t005
Table 6. Candidate gene.

| Gene_ID  | Gene_name | CCGG | CCWGG |
|----------|-----------|------|-------|
| Gene16474 | EGLN3     | NA   | Up    |
| Gene3969  | MUT       | NA   | Down  |
| Gene21049 | FOXL2     | Down | NA    |
| Gene23019 | PDIP1     | NA   | Up    |
| Gene22580 | SERPINA7  | Down | NA    |
| Gene11622 | TLX2      | NA   | Down  |
| Gene19027 | RHEBL1    | NA   | Down  |
| Gene6923  | SYP       | NA   | Down  |
| Gene466   | ECSCR     | NA   | UP    |
| Gene21566 | AQP11     | NA   | UP    |
| Gene18167 | PLEKHF2   | Up   | NA    |
| Gene12369 | RPL8      | Down | NA    |
| Gene21551 | NONO      | NA   | Down  |
| Gene5891  | MPHOSPH6  | NA   | UP    |
| Gene6309  | KIF20B    | NA   | Up    |
| Gene16559 | PNO1      | UP   | NA    |
| Gene10955 | HCST      | Down | NA    |
| Gene14033 | PDLIM3    | NA   | Up    |
| Gene5702  | HOXB7     | NA   | Down  |
| Gene21485 | ITM2A     | Down | NA    |
| Gene17170 | 43357     | NA   | Up    |
| Gene24124 | SMNDC1    | Up   | NA    |
| Gene22641 | MRPL20    | NA   | Down  |
| Gene7198  | INPP1     | NA   | Down  |
| Gene12913 | S100A1    | NA   | Up    |
| Gene6225  | FAM169A   | NA   | Up    |
| Gene17999 | TOP2B     | NA   | Up    |
| Gene16857 | PCTP      | NA   | Down  |
| Gene16714 | KLHL12    | NA   | Up    |
| Gene9820  | LMBRD1    | NA   | Up    |
| Gene8883  | ASPM      | NA   | Up    |
| Gene4484  | RGCC      | NA   | Up    |
| Gene23671 | CCL19     | Down | NA    |
| Gene13542 | MED23     | NA   | Up    |
| Gene4051  | CDK5      | NA   | Down  |
| Gene17223 | LRRK2     | NA   | Up    |
| Gene14545 | FRG1      | Down | NA    |
| Gene18760 | CPA3      | NA   | Down  |
| Gene11835 | C1QTNF4   | NA   | Down  |
| Gene24004 | LPL       | UP   | NA    |
| Gene24901 | UBQLN2    | NA   | Down  |
| Gene16856 | TMEM100   | NA   | Up    |

(Continued)
| Gene_ID  | Gene_name | CCGG | CCWGG |
|---------|-----------|------|-------|
| Gene13441 | PYGO1    | NA   | Down |
| Gene10329 | NDUFA9   | NA   | UP    |
| Gene1943  | TP53I3   | NA   | UP    |
| Gene10316 | BBOX1    | NA   | UP    |
| Gene17497 | IGSF6    | UP   | NA    |
| Gene20189 | UBE2C    | Down | NA    |
| Gene22225 | MED17    | NA   | UP    |
| Gene21475 | HOXD12   | UP   | NA    |
| Gene9863  | DHX15    | NA   | UP    |
| Gene19229 | NFATC2IP | NA   | Down |
| Gene23893 | ABT1     | NA   | UP    |
| Gene14704 | TF82M    | NA   | UP    |
| Gene8524  | LMOD2    | NA   | UP    |
| Gene12573 | PTMS     | NA   | Down |
| Gene3328  | ERP44    | NA   | UP    |
| Gene15160 | AP0C4    | Down | Down |
| Gene22228 | HEPHL1   | NA   | UP    |
| Gene21241 | SRPK3    | NA   | Down |
| Gene14716 | SOX9     | Down | NA    |
| Gene5821  | NKX2-5   | UP   | UP    |
| Gene2007  | PPRC1    | NA   | UP    |
| Gene12983 | PSMD10   | NA   | Down |
| Gene5916  | OPALIN   | NA   | UP    |
| Gene12986 | ATP6V1C1 | NA   | UP    |
| Gene16429 | EMC4     | UP   | NA    |
| Gene19144 | MSMO1    | NA   | Down |
| Gene16982 | SOX17    | NA   | Down |
| Gene24358 | MAGEH1   | Down | NA    |
| Gene19794 | SCGB3A1  | NA   | UP    |
| Gene11028 | WDR83    | NA   | Down |
| Gene3820  | PGM2L1   | NA   | Down |
| Gene11525 | FDX1L    | NA   | UP    |
| Gene10761 | GPR88    | UP   | NA    |
| Gene21912 | RPS7     | NA   | UP    |
| Gene7562  | GNAI1    | NA   | UP    |
| Gene8984  | SRD5A3   | NA   | Down |
| Gene17461 | SLC25A20 | NA   | UP    |
| Gene13012 | CLUL1    | NA   | UP    |
| Gene13916 | TUBD1    | NA   | UP    |
| Gene20231 | EDA2R    | NA   | Down |
| Gene13768 | ACVR2A   | NA   | UP    |
| Gene14993 | CLDN17   | UP   | NA    |
| Gene6350  | BTLA     | NA   | UP    |
| Gene703   | TMEM158   | NA   | Down |
| Gene17409 | LUC7L3   | NA   | UP    |
| Gene21871 | ADAM2    | NA   | UP    |
| Gene14058 | HOXC6    | NA   | Down |

(Continued)
functions of p53-mediated signaling pathway and promote the development of cataract. In addition, other direct evidence comes from the certain genes that have been previously reported to be associated with cataract pathogenesis. For example, Glutathione S-Transferase Mu 3 (GSTM3, hypermethylated in giant pandas with cataract) was considered to prevent the age-related cataract by protecting the lens from oxidative stress and a decreased expression level of GSTM3 was observed in the lens tissue of patients with age-related cataract, which correlated with the hypermethylation of GSTM3 promoters[32].

Since DNA methylation is reversible and can be influenced by the external factors[33], the research on the appropriate epigenetic drugs based on the specific cataract-associated genes would be wildly used in prevention of age-related cataract development in giant pandas. In this study, we shed light on the methylation characteristics of giant panda suffering from cataract and provide a number of candidate epigenetic therapeutic targets for the prevention and treatment of cataract in the aged giant panda. Nevertheless, the small sample size and the lack of functional experiments limit the practical utility of the findings in this study. Therefore, further efforts are needed to address the issues as follows: 1) the validation of DMGs in a large giant panda population; 2) the influences of certain aberrant DNA methylation events on gene activity; 3) the key genes with major contributions to the cataract development in age giant pandas; 4) the molecular mechanisms of key genes in the pathogenesis of age-related cataract. In addition, we also observed that some giant pandas with cataract can be self-healing after their living environments were changed. The contribution of reversible epigenetic modifications (e.g., DNA methylation) caused by environmental stimulus to this phenomenon will be explored in our future studies.

## Conclusion

In short, we determined a number of DMGs that had potential roles in regulating the activity of cataract-related pathways, such as base excision repair, apoptosis, and p53 signaling pathway. Moreover, these findings were further supported by detailed genes with abnormal methylation pattern in giant pandas with cataract. For example, the CASP3 gene encodes a cysteine-aspartic acid protease that served to as an apoptosis executor, and has been linked with cataract
HMGB1 plays an important role in protecting the keratinocytes from ultraviolet radiation-induced cell death and is thus involved in cataract formation[36]. Overall, all the results argue for an important role of aberrant methylation changes in the development of cataract in aged giant pandas.

**Supporting information**

S1 Table. Methylation level of gene region.  
(XLSX)

S2 Table. Genes associated with different methylation levels.  
(XLS)

S3 Table. KEGG enrichment of genes related to different methylation levels.  
(XLSX)

S4 Table. GO enrichment of genes related to different methylation levels.  
(XLSX)

S5 Table. Cataract related genes have been reported.  
(XLS)

**Author Contributions**

**Conceptualization:** Yuyan You, Xuefeng Liu.

**Data curation:** Yuyan You.

**Formal analysis:** Yuyan You.

**Funding acquisition:** Yuyan You.

**Investigation:** Yuyan You.

**Methodology:** Chao Bai.

**Resources:** Yuyan You, Xuefeng Liu, Maohua Xia, Ting Jia, Xiaoguang Li, Chenglin Zhang, Yucun Chen, Sufen Zhao, Liqin Wang, Wei Wang, Yanqiang Yin, Yunfang Xiu, Lili Niu, Jun Zhou, Tao Ma, Yang Du, Yanhui Liu.

**Software:** Yuyan You.

**Visualization:** Yuyan You.

**Writing – original draft:** Yuyan You, Chao Bai.

**References**

1. Uno HJA. Age-related pathology and biosenescence markers in captive rhesus macaques. Age. 1997; 20:1–13. https://doi.org/10.1007/s11357-997-0001-5 PMID: 23604287

2. Asbell PA, Dualan I, Mindel JS, Brocks D, Ahmad M, Epstein SPJTL. Age-related cataract. The Lancet. 2005; 365:599–609.

3. Urfer SR, Greer K, Wolf NS. Age-related cataract in dogs: a biomarker for life span and its relation to body size. Age. 2011; 33:451–60. https://doi.org/10.1007/s11357-010-9158-4 PMID: 20607428

4. Truscott RJWJEER. Age-related nuclear cataract-oxidation is the key. Experimental Eye Research. 2005; 80:709–25. https://doi.org/10.1016/j.exer.2004.12.007 PMID: 15862178

5. Jin Y, Lin W, Huang S, Zhang C, Pu T, Ma W, et al. Dental Abnormalities in Eight Captive Giant Pandas (*Ailuropoda melanoleuca*) in China. Journal of Comparative Pathology. 2012; 146:357–64. https://doi.org/10.1016/j.jcpa.2011.08.001 PMID: 21906751
6. Jin Y, Chen S, Chao Y, Pu T, Xu H, Liu X, et al. Dental Abnormalities of Eight Wild Qinling Giant Pandas (Ailuropoda Melanoleuca Qinlingensis), Shaanxi Province, China. Journal of Wildlife Disease. 2015; 51:849–59.

7. Hammond CJ, Duncan DD, Snieder H, De Lange M, West SK, Spector TD, et al. The heritability of age-related cortical cataract: The twin eye study. Investigative Ophthalmology & Visual Science. 2001; 42:601–5.

8. Ottonello S, Foroni C, Carta A, Petrucco S, Maraini GJO. Oxidative Stress and Age-Related Cataract. Ophthalmologica. 2000; 214:78–85. https://doi.org/10.1159/000027474 PMID: 10657746

9. Ho M, Peng Y, Chen S, Chioo SJJoCG, Geriatrics. Senile cataracts and oxidative stress. Journal of Clinical Gerontology and Geriatrics. 2010; 1:17–21.

10. Tinaztepe OE, Ay M, Esen EJCER. Nuclear and Mitochondrial DNA of Age-Related Cataract Patients Are Susceptible to Oxidative Damage. Current Eye Research. 2017; 42:1–6. https://doi.org/10.1080/02713683.2016.1175019

11. Billingsley G, Santhiya ST, Paterson AD, Ogata K, Wodak SJ, Hosseini SM, et al. CRYBA4, a Novel Human Cataract Gene, Is Also Involved in Microphthalmia. American Journal of Human Genetics. 2006; 79:702–9. https://doi.org/10.1086/507712 PMID: 16960806

12. Hasanova N, Kubo E, Kumamoto Y, Takamura Y, Akagi YJBJoO. Age-related cataracts and Prdx6: correlation between severity of lens opacity, age and the level of Prdx 6 expression. British Journal of Ophthalmology. 2009; 93:1081–4. https://doi.org/10.1136/bjo.2008.152272 PMID: 19429582

13. Zhang Y, Zhang L, Sun D, Li Z, Wang L, Liu PJMV. Genetic polymorphisms of superoxide dismutases, catalase, and glutathione peroxidase in age-related cataract. Molecular Vision. 2011; 17:2325–32. PMID: 21921984

14. Kim JK, Samaranayake M, Pradhan SJC, Sciences ML. Epigenetic mechanisms in mammals. Cellular and Molecular Life Sciences. 2009; 66:596–612. https://doi.org/10.1007/s00018-008-8432-4 PMID: 18985277

15. Xiao F, Kong Q, Perry B, He Y. Progress on the role of DNA methylation in aging and longevity. Briefings in Functional Genomics. 2016; 15:454–9. https://doi.org/10.1093/bfgp/elw009 PMID: 27032421

16. Liu X, Luo Y, Zhou P, Lu YJO, Science V. DNA methylation mediated and oxidative stress related genes CRYAA and GJA3 in nuclear age-related cataract (ARC) and its mechanism. Investigative Ophthalmology & Visual Science. 2015; 56:5877.

17. Zhou P, Luo Y, Liu X, Fan L, Lu YJTFJ. Down-regulation and CpG island hypermethylation of CRYAA in age-related nuclear cataract. The FASEB Journal. 2012; 26:4897–902. https://doi.org/10.1096/fj.12-213702 PMID: 22889833

18. Wang Y, Li F, Zhang G, Kang L, Qin B, Guan HJCER. Altered DNA Methylation and Expression Profiles of 8-Oxoguanine DNA Glycosylase 1 in Lens Tissue from Age-related Cataract Patients. Current Eye Research. 2015; 40:815–21. https://doi.org/10.3109/02713683.2014.957778 PMID: 25310012

19. Wang S, Lv J, Zhang L, Dou J, Sun Y, Li X, et al. MethylRAD: a simple and scalable method for genome-wide DNA methylation profiling using methylation-dependent restriction enzymes. Open Biology. 2015; 5:150130. https://doi.org/10.1098/rsob.150130 PMID: 26581575

20. Li R, Yu C, Li Y, Lam TW, Yu S, Kristiansen K, et al. SOAP2: an improved ultrafast tool for short read alignment. Bioinformatics. 2009; 25:1966–7. https://doi.org/10.1093/bioinformatics/btp336 PMID: 19497933

21. Shu X, Shu S, Cheng H, Tang S, Yang L, Li H, et al. Genome-Wide DNA Methylation Analysis During Palatal Fusion Reveals the Potential Mechanism of Enhancer Methylation Regulating Epithelial Mesenchyme Transformation. DNA and Cell Biology. 2018; 37:560–73. https://doi.org/10.1089/dna.2018.4141 PMID: 29608334

22. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics. 2010; 26:139–40. https://doi.org/10.1093/bioinformatics/btp166 PMID: 19910308

23. Cingolani P, Platts AE, Wang LL, Coon M, Nguyen T, Wang L, et al. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff. Fly. 2012; 6:80–92. https://doi.org/10.4161/fly.19695 PMID: 22728672

24. Quinlan AR, Hall IMJB. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics. 2010; 26:841–2. https://doi.org/10.1093/bioinformatics/btp033 PMID: 2101278

25. Kanehisa M, Goto SJNAR. KEGG: Kyoto Encyclopedia of Genes and Genomes. Nucleic Acids Res. 1999; 28:27–30.

26. Ashburner M, Ball CA, Blake JA, Botstein D, Butler HL, Cherry JM, et al. Gene ontology: tool for the unification of biology. Nature Genetics. 2000; 25:25–9. https://doi.org/10.1038/75556 PMID: 10802651
27. Davis AP, King BL, Mockus S, Murphy CG, Saracenirichards CA, Rosenstein MT, et al. The Comparative Toxicogenomics Database: update 2011. Nucleic Acids Res. 2011; 39:1067–72.

28. Li BY, Han JA, Im JS, Morrone A, Johung K, Goodwin EC, et al. Senescence-associated β-galactosidase is lysosomal β-galactosidase. Aging Cell. 2006; 5:187–95. https://doi.org/10.1111/j.1474-9726.2006.00199.x PMID: 16626397

29. Lawless C, Wang C, Jurk D, Merz A, Zglinicki T, Passos J. Quantitative assessment of markers for cell senescence. Experimental Gerontology. 2010; 45:772–8. https://doi.org/10.1016/j.exger.2010.01.018 PMID: 20117203

30. Charakidas A, Kalogeraki A, Tsilimbaris MK, Koukoulomatis P, Brouzas D, Delides G. Lens epithelial apoptosis and cell proliferation in human age-related cortical cataract. European Journal of Ophthalmology. 2005; 15:213–20. https://doi.org/10.1177/112067210501500206 PMID: 15812762

31. Ji W, Science V. αA-Crystallin Regulates p53-Mediated Signaling Pathway to Prevent Apoptosis of Lens Epithelial Cells and Cataractogenesis. Investigative Ophthalmology & Visual Science. 2012; 53:1043–1043.

32. Li B, Zhou J, Zhang G, Wang Y, Kang L, Wu J, et al. Relationship Between the Altered Expression and Epigenetics of GSTM3 and Age-Related Cataract. Investigative Ophthalmology & Visual Science. 2016; 57:4721–32.

33. Dor Y, Cedar H. Principles of DNA methylation and their implications for biology and medicine. The Lancet. 2018; 392:777–786.

34. Doshna CW, Fortner JH, Pfohl JC, Aleo ME, Verdugo M. Investigation of the Role of Apoptosis in Drug-induced Cataract Formation. Investigative Ophthalmology & Visual Science. 2002; 43:2377.

35. Galichanin K, Svedlund J, Soderberg P. Kinetics of GADD45α, TP53 and CASP3 gene expression in the rat lens in response to exposure to double threshold dose of UV-B radiation. Experimental Eye Research. 2012; 97:19–23. https://doi.org/10.1016/j.exer.2012.02.003 PMID: 22559303

36. Mou K, Liu W, Han D, LiPJ. HMGB1/RAGE axis promotes autophagy and protects keratinocytes from ultraviolet radiation-induced cell death. Journal of Dermatological Science. 2017; 85:162–9. https://doi.org/10.1016/j.jdermsci.2016.12.011 PMID: 28012822