

# SCIENTIFIC REPORTS



OPEN

## Hyperlipidemia-associated gene variations and expression patterns revealed by whole-genome and transcriptome sequencing of rabbit models

Received: 25 February 2016

Accepted: 11 May 2016

Published: 01 June 2016

Zhen Wang<sup>1,\*</sup>, Jifeng Zhang<sup>2,\*</sup>, Hong Li<sup>1,\*</sup>, Junyi Li<sup>1,\*</sup>, Manabu Niimi<sup>3</sup>, Guohui Ding<sup>1,4</sup>, Haifeng Chen<sup>4,5</sup>, Jie Xu<sup>2</sup>, Hongjiu Zhang<sup>6</sup>, Ze Xu<sup>7</sup>, Yulin Dai<sup>8,9</sup>, Tuantuan Gui<sup>8,9</sup>, Shengdi Li<sup>8,9</sup>, Zhi Liu<sup>8,9</sup>, Sujuan Wu<sup>4,10</sup>, Mushui Cao<sup>4,11</sup>, Lu Zhou<sup>8,9</sup>, Xingyu Lu<sup>5</sup>, Junxia Wang<sup>5</sup>, Jing Yang<sup>4,10</sup>, Yunhe Fu<sup>8,9</sup>, Dongshan Yang<sup>2</sup>, Jun Song<sup>2</sup>, Tianqing Zhu<sup>2</sup>, Shen Li<sup>3</sup>, Bo Ning<sup>3</sup>, Ziyun Wang<sup>3</sup>, Tomonari Koike<sup>12</sup>, Masashi Shiomi<sup>12</sup>, Enqi Liu<sup>13,14</sup>, Luonan Chen<sup>8</sup>, Jianglin Fan<sup>3,14</sup>, Y. Eugene Chen<sup>2</sup> & Yixue Li<sup>1,4,5,10,11</sup>

The rabbit (*Oryctolagus cuniculus*) is an important experimental animal for studying human diseases, such as hypercholesterolemia and atherosclerosis. Despite this, genetic information and RNA expression profiling of laboratory rabbits are lacking. Here, we characterized the whole-genome variants of three breeds of the most popular experimental rabbits, New Zealand White (NZW), Japanese White (JW) and Watanabe heritable hyperlipidemic (WHHL) rabbits. Although the genetic diversity of WHHL rabbits was relatively low, they accumulated a large proportion of high-frequency deleterious mutations due to the small population size. Some of the deleterious mutations were associated with the pathophysiology of WHHL rabbits in addition to the *LDLR* deficiency. Furthermore, we conducted transcriptome sequencing of different organs of both WHHL and cholesterol-rich diet (Chol)-fed NZW rabbits. We found that gene expression profiles of the two rabbit models were essentially similar in the aorta, even though they exhibited different types of hypercholesterolemia. In contrast, Chol-fed rabbits, but not WHHL rabbits, exhibited pronounced inflammatory responses and abnormal lipid metabolism in the liver. These results provide valuable insights into identifying therapeutic targets of hypercholesterolemia and atherosclerosis with rabbit models.

<sup>1</sup>Key Lab of Computational Biology, CAS-MPG Partner Institute for Computational Biology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai, China. <sup>2</sup>Center for Advanced Models for Translational Sciences and Therapeutics, University of Michigan Medical Center, Ann Arbor, MI, USA. <sup>3</sup>Department of Molecular Pathology, Interdisciplinary Graduate School of Medicine and Engineering, University of Yamanashi, Yamanashi, Japan. <sup>4</sup>Shanghai Center for Bioinformation Technology, Shanghai Industrial Technology Institute, Shanghai, China. <sup>5</sup>School of Life Science and Biotechnology, Shanghai Jiaotong University, Shanghai, China. <sup>6</sup>Department of Computational Medicine and Bioinformatics, University of Michigan, Ann Arbor, MI, USA. <sup>7</sup>EG Information Technology Enterprise (EGI), BasePair Biotechnology Co., Ltd., Shanghai, China. <sup>8</sup>Key Lab of Systems Biology, Institute of Biochemistry and Cell Biology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai, China. <sup>9</sup>University of Chinese Academy of Sciences, Beijing, China. <sup>10</sup>School of Biotechnology, East China University of Science and Technology, Shanghai, China. <sup>11</sup>School of Life Science and Technology, Shanghai Tongji University, Shanghai, China. <sup>12</sup>Institute for Experimental Animals, Kobe University School of Medicine, Kobe, Japan. <sup>13</sup>Research Institute of Atherosclerotic Disease and Laboratory Animal Center, Xi'an Jiaotong University School of Medicine, Xi'an, China. <sup>14</sup>Department of Pathology, Xi'an Medical University, Xi'an, China. \*These authors contributed equally to this work. Correspondence and requests for materials should be addressed to J.F. (email: jianglin@yamanashi.ac.jp) or Y.E.C. (email: echenum@umich.edu) or Y.L. (email: yxli@sibs.ac.cn)

Breed	DNA					RNA			
	Sample size	Depth of coverage per sample	Nucleotide diversity $\pi$	Watterson $\theta$	Tajima $D$	Diet	Tissue collected	Sample size	Data (Gb) per tissue per sample
NZW	$n = 10$	13.34	$2.80 \times 10^{-3}$	$2.26 \times 10^{-3}$	0.77	Standard chow diet	Aorta, liver, heart, kidney	$n = 4$	5.10
						0.3% cholesterol diet	Aorta, liver, heart, kidney	$n = 4$	5.13
JW	$n = 10$	12.94	$1.61 \times 10^{-3}$	$1.37 \times 10^{-3}$	-0.01	Standard chow diet	Aorta, liver, heart, kidney	$n = 4$	5.93
WHHL	$n = 10$	12.75	$1.44 \times 10^{-3}$	$1.19 \times 10^{-3}$	-0.08	Standard chow diet	Aorta, liver, heart, kidney	$n = 4$	5.70

**Table 1. Experimental design and sequencing data.**

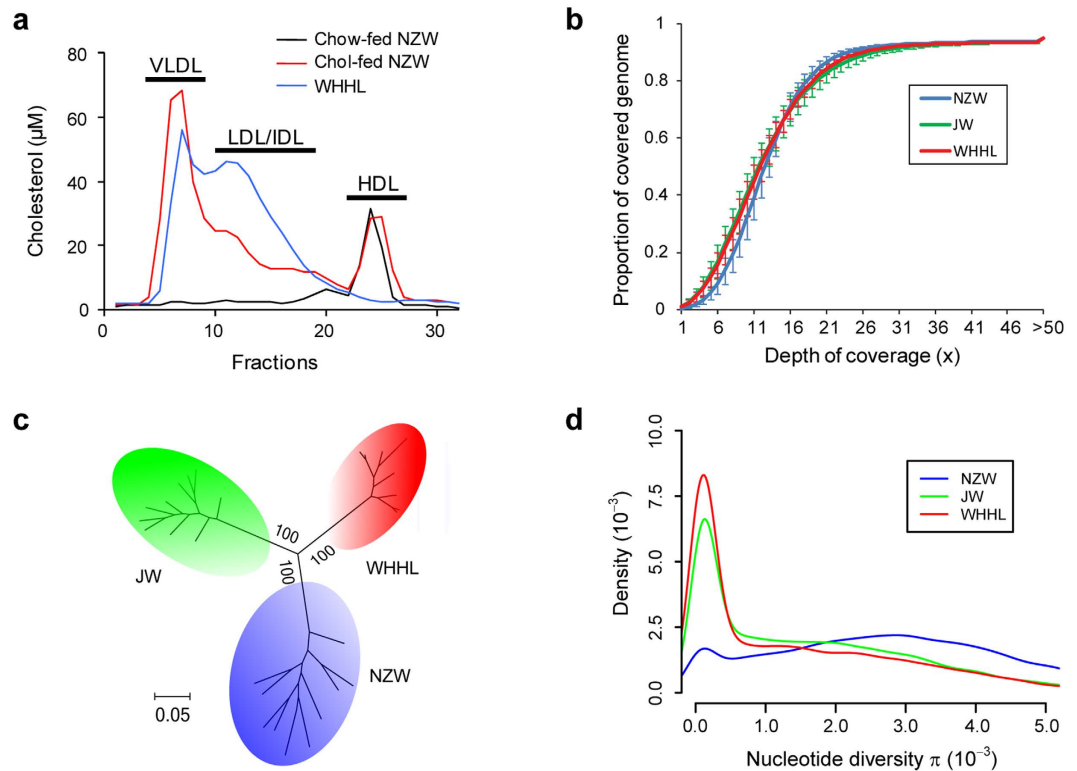
The European rabbit (*Oryctolagus cuniculus*) is an important experimental animal model for biomedical science. Rabbits are not only the most-used animal for the production of antibodies, but also they are widely used for studying a variety of human diseases, such as infectious disease, neoplasms, ophthalmic disease, Alzheimer's disease, and respiratory disease<sup>1</sup>. Like humans, but unlike rodents, such as mice and rats, rabbits have unique features of lipid metabolism that have made them an important model for human hyperlipidemia and atherosclerosis<sup>2</sup>. The first experiment using rabbits to investigate atherosclerosis was performed more than a century ago<sup>3</sup>. When fed a diet rich in cholesterol, laboratory rabbits rapidly develop hypercholesterolemia and atherosclerosis<sup>2</sup>. In addition, genetic defects in low density lipoprotein receptor (*LDLR*) in Watanabe heritable hyperlipidemic (WHHL) rabbits can lead to spontaneous hypercholesterolemia and atherosclerosis, even when they are on a normal chow diet<sup>4,5</sup>. Therefore, the rabbit model has provided tremendous breakthroughs and insights into understanding the molecular and cellular mechanisms of atherosclerosis, including the discoveries of *LDLR* deficiency as a cause for human familial hypercholesterolemia<sup>6</sup> and statin, the most potent lipid-lowering drug<sup>7</sup>, which is prescribed annually for more than 30 million hyperlipidemic patients worldwide<sup>8,9</sup>.

Despite the importance of rabbit models for the study of hypercholesterolemia and atherosclerosis, genomic and transcriptomic information related to hyperlipidemia and atherosclerosis is lacking, which hampers the use of rabbits for translational research<sup>2</sup>. Recently, a high-quality reference genome for the European rabbit with references to domestication and speciation was reported<sup>10,11</sup>. In the current study, we performed whole-genome sequencing on three breeds of popular experimental rabbits, wild-type New Zealand White (NZW), Japanese White (JW) and WHHL rabbits, in an attempt to identify whether there are other gene mutations or modifiers that may be involved in the pathogenesis of hypercholesterolemia and atherosclerosis in WHHL rabbits. In addition to WHHL rabbits, cholesterol-rich diet (Chol)-fed rabbits are often used as a model for the study of human hypercholesterolemia and atherosclerosis<sup>2</sup>. While both WHHL and Chol-fed rabbits exhibit hypercholesterolemia and atherosclerosis, the gene expression profiles of atherosclerotic lesions and livers have not been systemically investigated. Toward this goal, we conducted deep transcriptome sequencing of the aortas, livers, hearts and kidneys derived from the two hypercholesterolemic models along with wild-type control rabbits. These results provide valuable resources for the investigation of hypercholesterolemia and atherosclerosis using rabbit models.

## Results

**Whole-genome sequencing of laboratory rabbits.** We collected three common breeds of laboratory rabbits: NZW, JW and WHHL rabbits (Table 1). On a standard chow diet, both NZW and JW rabbits have relatively low plasma cholesterol levels compared to humans, and their cholesterol is mainly carried by high density lipoproteins (HDLs, Fig. 1a). WHHL rabbits are genetically deficient in *LDLR* function; thus, they develop hypercholesterolemia and atherosclerosis, even on a standard chow diet. Normal rabbits can also develop hypercholesterolemia and atherosclerosis when fed a diet rich in cholesterol (Chol). Although both WHHL and Chol-fed rabbits exhibit hypercholesterolemia, their lipoprotein profiles are quite different; WHHL rabbits have increased levels of LDL-cholesterol accompanied by low HDLs, while Chol-fed rabbits have increased hepatically and intestinally derived remnant lipoproteins, called  $\beta$ -VLDL (Fig. 1a).

We performed whole-genome sequencing of 10 rabbits for each of the three breeds (Supplementary Fig. S1), resulting in a depth of coverage of approximately  $13\times$  for each individual after alignment to the reference genome (Fig. 1b and Supplementary Fig. S2). Totally, we identified 29.8 million SNPs (Supplementary Fig. S3) and 1.6 million small indels (Supplementary Fig. S4) in the 30 genomes. Phylogenetic tree building (Fig. 1c) and principal component analysis (Supplementary Fig. S5) based on genome-wide SNPs conformed distinct genetic backgrounds of the three breeds. Most of the rabbits were assumed to be unrelated except two pairs of WHHL rabbits (Supplementary Fig. S5). The genetic diversity of NZW rabbits (Table 1), measured by the nucleotide diversity  $\pi$  (Fig. 1d) and Watterson's  $\theta$  (Supplementary Fig. S6), was consistent with a recent report<sup>11</sup>. It was also higher than that of JW and WHHL rabbits, suggesting that even though all of the breeds originated from European rabbits, NZW rabbits are derived from a larger population of progenitors. Furthermore, the Tajima's  $D$  (Table 1 and Supplementary Fig. S6) of NZW rabbits was positive and the largest among the three breeds, suggesting a moderate population bottleneck (sharp reduction in population size) during domestication<sup>12</sup>. In contrast, the Tajima's  $D$  of WHHL rabbits was negative and the smallest value, which is consistent with the fact that the breed

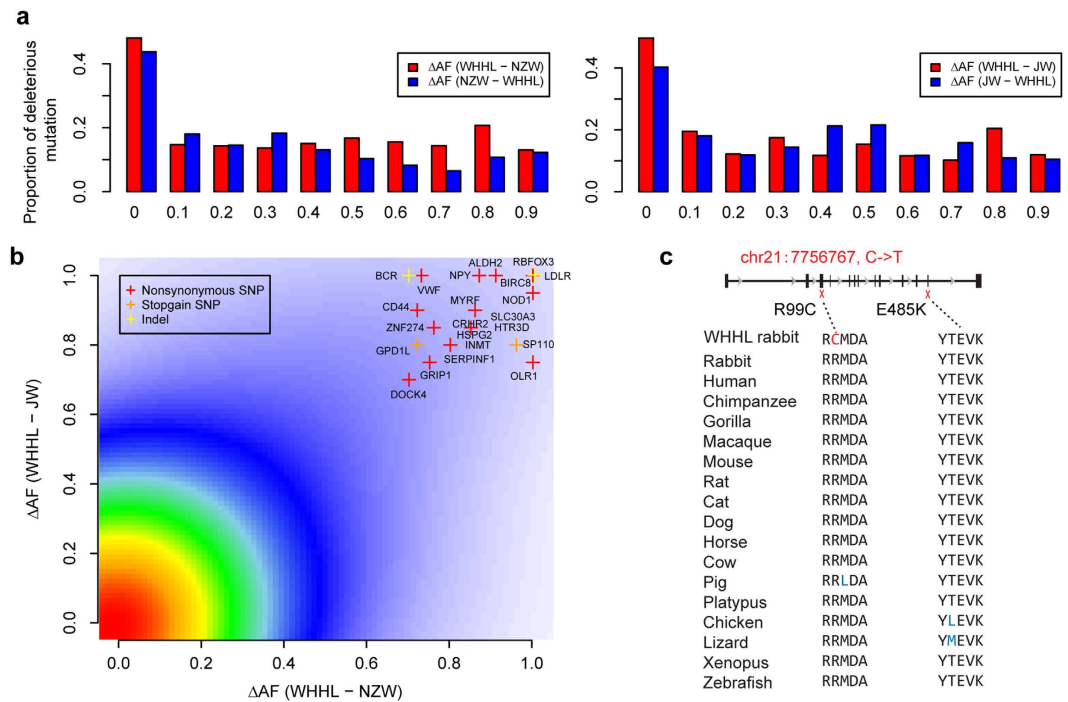


**Figure 1. Whole-genome sequencing of laboratory rabbits.** (a) Lipid profiles of standard chow-fed NZW, Chol-fed NZW and WHHL rabbits analyzed by high performance liquid chromatography. The Chol-fed NZW showed elevated  $\beta$ -VLDLs, and the WHHL rabbits showed increased LDLs and reduced HDLs. (b) Cumulative distribution of depth of coverage for whole-genome sequencing. The average depth of coverage was  $13\times$  for each individual rabbit. (c) Phylogenetic tree of the rabbits. The tree was constructed on the basis of representative SNPs with the maximum likelihood method. Bootstrap values are marked on the branch. (d) Distribution of nucleotide diversity  $\pi$ . The statistics were calculated for every 100 kb sliding-window across the genome.

underwent a severe population bottleneck during artificial selection. The level of linkage disequilibrium was lowest in NZW rabbits and highest in WHHL rabbits (Supplementary Fig. S7), which is also in agreement with their breeding history.

**Deleterious mutations in WHHL rabbits.** Although WHHL rabbits are well-known for their *LDLR* mutation as a cause of hypercholesterolemia<sup>5</sup>, it is possible that other deleterious mutations could rise to high frequency by genetic drift due to the extremely small population size of the breed. To search for such deleterious mutations possibly involved in cardiovascular diseases, we compiled a comprehensive gene list associated with cardiovascular diseases from both the knowledge database and human genome-wide association studies (Supplementary Fig. S8). Based on the functional annotations of SNPs and indels, we predicted deleterious mutations in the prior genes using the following criteria: 1) the mutation should alter the protein sequence; 2) for non-synonymous SNPs, it should have a SIFT score  $<0.05$ <sup>13</sup>. We examined the proportion of deleterious mutations by the difference in allele frequency ( $\Delta AF$ ) between WHHL and NZW or JW rabbits (Fig. 2a). It was shown that when  $\Delta AF > 0.4$  WHHL rabbits harbor a larger proportion of deleterious mutations with high frequency than NZW rabbits, which supported our hypothesis that deleterious mutations were easier to accumulate in the WHHL breed. As the genetic diversity and hence the population size of JW rabbits were quite close to WHHL rabbits (Table 1), this result could only be observed when  $\Delta AF > 0.8$  between the two breeds.

We focused on the putative deleterious mutations specifically enriched in WHHL rabbits with a criterion of  $\Delta AF > 0.7$  compared with both normal breeds, which resulted in 24 deleterious mutations in the prior genes in addition to the known 12-bp in-frame deletion in *LDLR*<sup>5</sup> (Fig. 2b and Supplementary Table S1). One of the deleterious mutations was located in *ALDH2*, the activation of which was found to decrease aortic atherosclerosis<sup>14</sup>. It is well-known that a loss-of-function mutation in human *ALDH2*, E487K, causes alcohol flushing and is associated with an increased risk of cardiovascular diseases<sup>15</sup>. Similar to this mutation in humans, the deleterious mutation that we predicted in the WHHL rabbits, R99C, occurred at a conserved site, which was invariant among all the vertebrates we examined (Fig. 2c and Supplementary Fig. S9). The mutant allele frequency remained low in the normal rabbits (NZW 10%, JW 0%) but became fixed in WHHL rabbits (100%). Another putative deleterious mutation was located in *VWF* (Fig. 2b), the mediator of blood coagulation. A mutation at the same position in the human *VWF* protein was reported in patients with type I von Willebrand's disease<sup>16</sup>. As a hypercoagulable state was found in WHHL rabbits<sup>17</sup>, the role of the deleterious mutation of *VWF* deserves further investigation. There

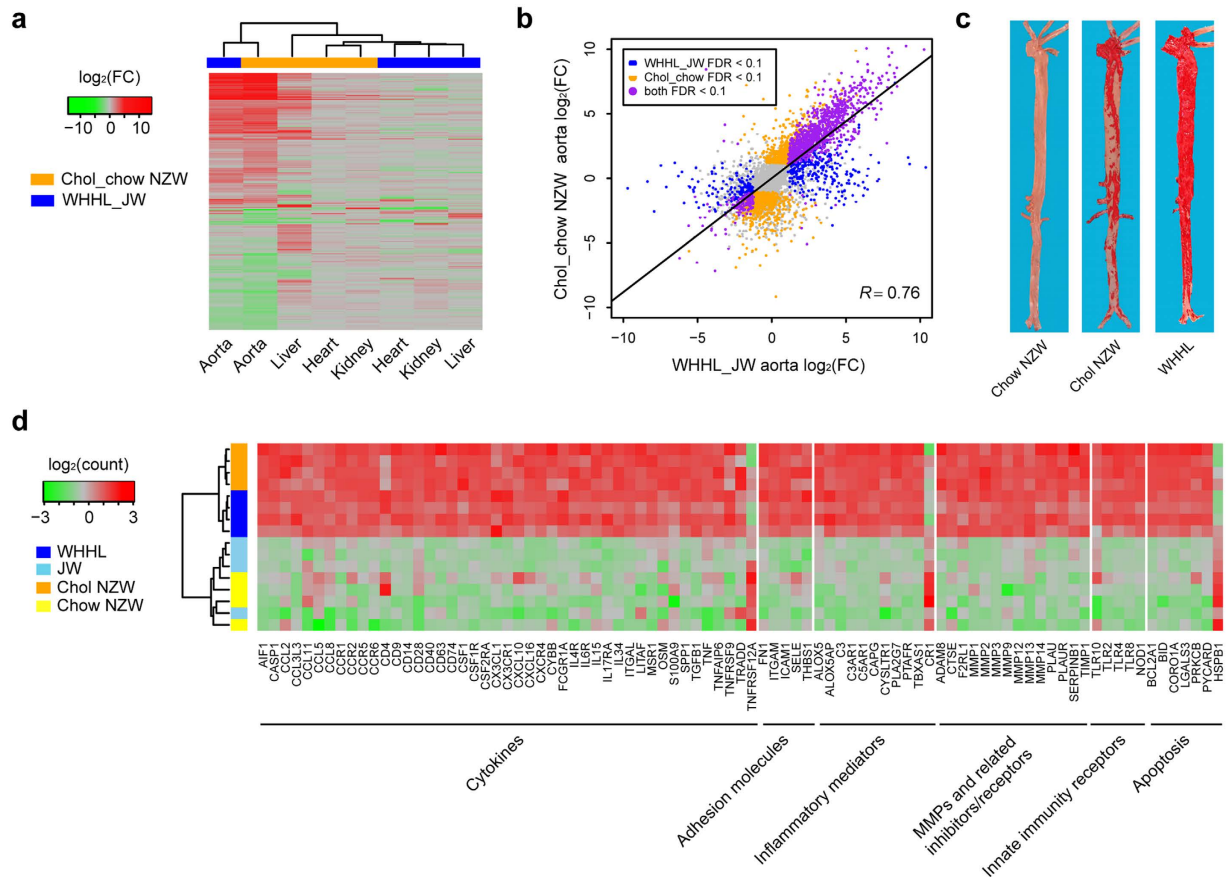


**Figure 2. Deleterious mutations in WHHL rabbits.** (a) Proportion of deleterious mutation by difference in allele frequency ( $\Delta$ AF) between WHHL and NZW or JW rabbits. (b) Putative deleterious mutations in WHHL rabbits with  $\Delta$ AF > 0.7 compared with both normal rabbits. Colors show the density of SNPs from high (red) to low (blue). Genes harboring deleterious mutations are highlighted. (c) Non-synonymous mutations in *ALDH2*. The red cross indicates the locations of mutations. R99C is a putative deleterious mutation in WHHL rabbits. E487K is a known loss-of-function mutation in humans. Both mutations occur at highly conserved sites across vertebrates.

was also a deleterious mutation identified in *OLRI* (Fig. 2b), and the gene was highly expressed in the atherosclerotic lesions of WHHL rabbits<sup>18</sup>. These results suggested that the deleterious mutations could function as genetic modifiers in the pathophysiology of WHHL rabbits.

**Transcriptome profiling of WHHL and Chol-fed rabbits.** We conducted deep transcriptome sequencing of different organs collected from both normal and hypercholesterolemic rabbits (Table 1 and Supplementary Figs S10 and 11). More than 77% of the assembled transcripts were isoforms of 15,760 known genes recorded in Ensembl (v76)<sup>19</sup> (Supplementary Fig. S12). Samples originating from the same tissue were clustered together based on their gene expression profiles (Supplementary Fig. S13). We performed differential expression analysis for Chol-fed versus normal chow-fed NZW rabbits and WHHL versus JW rabbits (Supplementary Figs S14 and 15). The results showed that differentially expressed genes (DEGs) predominantly occurred in the aortas of both hypercholesterolemic models (2,719 and 1,627, respectively, false discovery rate [FDR] < 0.1). In the liver, however, the transcriptional changes were notable in the Chol-fed rabbits (1,403 DEGs) but not in the WHHL rabbits (181 DEGs). The expression changes in the heart and kidney in both Chol-fed and WHHL rabbits were not remarkable compared to each wild-type rabbit.

For a more detailed model-by-tissue comparison, we selected DEGs observed in at least one condition and performed clustering of the conditions based on the  $\log_2$ -fold change ( $\log_2$ FC) of the DEGs (Fig. 3a). The result illustrates similar transcriptional responses in the aortas of Chol-fed and WHHL rabbits but different transcriptional responses in the livers of the two models. The magnitudes of aortic changes were highly positively correlated in the two models ( $R = 0.76$ , Fig. 3b). This result is in agreement with the pathological changes in the aortic lesions of atherosclerosis that were grossly present in both Chol-fed and WHHL rabbits (Fig. 3c). To gain functional insights into the DEGs, we performed a comparison using Ingenuity Pathway Analysis (IPA). Both function and pathway enrichment analyses supported the similarity of aortic expression changes in Chol-fed and WHHL rabbits (Supplementary Fig. S16). Notably, inflammatory responses in the lesions were significantly activated in both models (activation  $z$ -score = 6.28 and 6.75, respectively), which is consistent with the notion that atherosclerosis is a chronic inflammatory disease initiated by lipid deposition in the arterial wall<sup>20</sup>. A common set of DEGs and pathways (Supplementary Fig. S17) responsible for inflammation was observed in multiple processes of lesion development, including cytokines and chemokines along with their receptors, which are involved in monocyte adhesion, activation, and migration as well as matrix metalloproteinases (MMPs), which are responsible for extracellular matrix degradation (Fig. 3d).



**Figure 3. Transcriptome profiling of rabbit models with aortic atherosclerosis.** (a) Heatmap of DEGs between Chol-fed and normal chow-fed NZW rabbits as well as between WHHL and JW rabbits. The DEGs should have FDR < 0.1 in at least one comparison. Log<sub>2</sub>-fold changes of DEGs are illustrated by gradient colors. The transcriptional changes of Chol-fed and WHHL rabbits were similar in the aorta but distinct in the liver. (b) Strong positive correlation of expression changes in the aorta between Chol-fed and WHHL rabbits. The correlation coefficient was calculated for DEGs in at least one condition. (c) Macrographs of aortas in normal chow-fed, Chol-fed and WHHL rabbits. The aortic lesions are stained red with Sudan IV. Both Chol-fed and WHHL rabbits showed extensive atherosclerotic lesions. (d) Heatmap of representative DEGs responsible for inflammation responses in the aorta. The read counts were log-transformed and normalized across samples. These genes induced inflammatory responses in both Chol-fed and WHHL rabbits compared with the normal controls.

**Hepatic transcriptome profiles of Chol-fed and WHHL rabbits.** Because the liver is an important organ to mediate lipid metabolism, we further investigated transcriptional changes in the livers of both Chol-fed and WHHL rabbits. As shown in Fig. 3a, both the number of DEGs and their expression patterns were dramatically changed in Chol-fed rabbits but not in WHHL rabbits. In fact, the magnitudes of the transcriptional changes in the liver displayed little correlation between the two models ( $R = -0.09$ , Fig. 4a). Functional analysis by IPA revealed that inflammatory responses were significantly activated in the livers of Chol-fed rabbits (activation z-score = 4.69). Furthermore, the DEGs mediating inflammatory responses were identified in the Chol-fed but not in the WHHL rabbits (Fig. 4b). Many inflammation-related pathways involved in aortic atherosclerosis were also invoked in the livers of Chol-fed rabbits (Supplementary Fig. S18), suggesting that cholesterol feeding can lead to hepatic injury and inflammatory changes.

To reveal the underlying metabolic impairment induced by a cholesterol-rich diet, we focused on the transcriptional control of hepatic lipid metabolism. IPA can explore upstream regulators through enrichment of their downstream DEGs, even if the regulators are not differentially expressed. This analysis identified several changes in key regulators related to cholesterol and fatty acid homeostasis in the livers of Chol-fed rabbits, including liver X receptors (*NR1H2* and *NR1H3*), farnesoid X receptor (*NR1H4*), *SREBF1*, *SREBF2*, *PPARA* and *PPARG*<sup>21</sup> (Fig. 4c and Supplementary Fig. S18). Liver X receptors are cholesterol-activated nuclear receptors, which target genes involved in reverse transport of excess cholesterol from peripheral tissues to the liver<sup>22</sup>. *ABCA1* (FDR =  $5.87 \times 10^{-8}$ ), *ABCG1* (FDR =  $9.74 \times 10^{-5}$ ), *PLTP* (FDR = 0.05) and *LPL* (FDR =  $3.89 \times 10^{-12}$ ), the target genes responsible for cholesterol transport and lipoprotein remodeling, were significantly up-regulated in the livers of Chol-fed rabbits (Supplementary Fig. S19). Another target gene, *MYLIP*, (FDR = 0.02) was also significantly overexpressed, which could trigger the ubiquitylation and degradation of LDLR<sup>22</sup>. *SREBF2* is a transcriptional factor responsible for cholesterol synthesis and is activated by low cellular cholesterol levels<sup>23</sup>. Not



deleterious mutations could also play a role in the pathogenesis of hypercholesterolemia and atherosclerosis in WHHL rabbits. For example, a high-frequency deleterious mutation was predicted at a quite conserved site on *ALDH2*, which was previously reported to play a cardio-protective role via both the elimination of toxic aldehydes under oxidative stress and the bioactivation of nitroglycerin to nitric oxide<sup>25</sup>. Recent studies found that Alda-1, an activator of *ALDH2*, can reduce myocardial infarct size induced by acute ischemia<sup>26</sup>. Administration of Alda-1 decreases aortic atherosclerosis in *Apoe* knockout mice<sup>14</sup>. Because WHHL rabbits suffer from hypercholesterolemia, atherosclerosis, myocardial infarction<sup>27</sup>, insulin resistance<sup>28</sup> and visceral adipose accumulation<sup>29</sup>, further studies are required to elucidate the pathophysiological significance of the deleterious mutations.

Although rabbits are frequently used to study human atherosclerosis, gene expression profiling of aortic and liver lesions has not been systemically elucidated. We conducted transcriptome sequencing of these two tissues of WHHL and Chol-fed NZW rabbits and compared them with those of control rabbits. Quantification of the transcriptome allowed us to identify more than 2,000 DEGs in aortic atherosclerotic lesions of WHHL and Chol-fed rabbits, as well as more than 1,000 DEGs in the livers of Chol-fed rabbits. We found that many inflammation-related genes were upregulated in aortic atherosclerotic lesions of both WHHL and Chol-fed rabbits, which is consistent with the notion that atherosclerosis is essentially evoked by inflammation<sup>30</sup>. In this regard, many cytokines and chemokines, along with their receptors and apoptosis signaling pathways, were markedly upregulated. In addition, many MMPs were upregulated in the lesions, suggesting that they are involved in lesion formation<sup>31</sup>. Nevertheless, gene expression profiles of the lesions were similar between WHHL and Chol-fed rabbits, suggesting that regardless of the different types of hypercholesterolemia in WHHL (predominated by LDLs) and Chol-fed (predominated by  $\beta$ -VLDL or remnant lipoproteins) rabbits, the process of lesion development occurs in a similar pathway, namely, inflammation initiated by lipid deposition in the intima. However, it should be noted that Chol-fed rabbits, but not WHHL rabbits, also exhibited a pronounced inflammatory response and abnormal lipid metabolism in the liver, suggesting that exogenous dietary cholesterol feeding not only elevates the plasma levels of cholesterol but also induces hepatic dysfunction. In fact, it has been reported that feeding rabbits with 1% cholesterol for 8 weeks can induce non-alcoholic fatty liver disease<sup>32</sup>.

In conclusion, we have successfully characterized the whole-genome variants and transcriptome profiles of the three breeds of laboratory rabbits. In a future study, we will investigate whether these deleterious mutations and DEGs play any roles in the pathogenesis of various disorders in the hyperlipidemic rabbits. These resources will be vital in identifying therapeutic targets for the treatment of hyperlipidemia and atherosclerosis.

## Methods

**Sample collection.** NZW rabbits were purchased from Covance Com., USA, and JW rabbits were purchased from Japan SLC Inc. (Shizuoka, Japan) and WHHL rabbits were provided by Kobe University School of Medicine, Japan. The rabbits were fed a normal chow diet. Animals were euthanized by overdose injection of pentobarbital solution. We collected liver tissue for extracting DNA for genome sequencing ( $n = 10$ ). For the transcriptome profiling, we also used NZW rabbits fed with a cholesterol-rich diet containing 0.3% cholesterol and 3% soybean oil for 16 weeks, which is commonly used for studying atherosclerosis. The liver, kidney, heart and aortic arch were collected for extracting total RNA for transcriptomic analysis ( $n = 4$ ). All animal experiments were performed with the approval of the Animal Care and Use Committee of the University of Yamanashi, Kobe University and University of Michigan. The methods were carried out in accordance with relevant and approved guidelines and regulations.

**Genome and RNA Sequencing.** Genomic DNA was extracted from 200  $\mu$ l liver tissue lysate with the QIAamp Blood DNA mini kit (Qiagen, Germany) according to the manufacturer's manual. DNA quality and integrity were controlled by the A260/280 ratio, agarose gel electrophoresis using Qubit 2.0 (Life Technologies, USA) and a Bioanalyzer 2100 (Agilent, Germany). For sequencing the library preparation, 1  $\mu$ g of genomic DNA was sheared to fragments of 300–400 bp, end-repaired, A-tailed and ligated to Illumina sequencing adapters. The ligated products were selected for a length of 400–500 bp on a 2% agarose gel and amplified by LM-PCR. The library was sequenced on an Illumina HiSeq2000 with 2  $\times$  100 bp paired-end mode, which was controlled by HiSeq Control Software. Total RNA sample quality and integrity were controlled with a Bioanalyzer 2100 (Agilent, Germany). For the library preparation, 3  $\mu$ g total RNA was captured by Dynabeads Oligo (Life Technologies, USA), sheared to fragments of 200 bp, and reverse transcribed by the SuperScriptIII cDNA Synthesis Kit (Life Technologies, USA). cDNA was end-repaired, A-tailed and ligated to Illumina sequencing adapters and amplified by PCR. Library preparation was performed with the TruSeq RNA LT V2 Sample Prep Kit (Illumina, San Diego). The sequencing library was qualified by Qubit 2.0 (Life Technologies, USA) and a Bioanalyzer 2100 (Agilent, Germany) and then sequenced on an Illumina HiSeq2000 with 2  $\times$  100 bp paired-end mode, which was controlled by HiSeq Control Software.

**DNA variant calling.** Raw sequencing reads were processed by in-house scripts to filter low quality reads and trim low quality bases (quality <20) at the 3' end of reads. Remaining reads were aligned to the rabbit reference genome, *oryCun2* by BWA-MEM (v0.7.4)<sup>33</sup>, with default parameters. Aligned reads were sorted and summarized by SAMtools (v0.1.9)<sup>34</sup>, where duplicated reads were removed. SNPs and genotypes were called by the SAMtools pipeline<sup>34</sup> with recommended parameters. The rabbits belonging to the same breed were combined for variant calling. SNPs with a depth of coverage lower than 20 or higher than 300 were filtered. Small indels were called by Dindel (v1.0.1)<sup>35</sup>, which performed local realignment around candidate indels. The recommended workflow for diploid samples was adopted. Only indels occurring in at least two samples were selected for realignment, and the maximum number of reads in a realignment window was 50. Variants of different rabbits were merged by VCFtools (v0.1.11)<sup>36</sup>. Functional effects of the variants after filtering were annotated by ANNOVAR

(v2013-06-21)<sup>37</sup>, using Ensembl (v76)<sup>19</sup> genes as the database. Predictions of non-synonymous SNPs on protein function were downloaded from the SIFT database for the rabbit genome (<http://sift-db.bii.a-star.edu.sg/>, v2014-10-07)<sup>13</sup>.

**Population genetics analysis.** Genetic relationships between all pairwise rabbits were estimated by KING<sup>38</sup>, which reported the proportion of SNPs with zero identical-by-state and calculated kinship coefficients. PCA clustering was performed by GCTA (v1.13)<sup>39</sup>. The phylogenetic tree was constructed by SNPhylo (v2014-07-01)<sup>40</sup>, which selected representative SNPs based on LD blocks and produced maximum likelihood trees. Standard population genetic statistics, including nucleotide diversity  $\pi$ , Watterson's  $\theta$  and Tajima's  $D$  were calculated using packages in Bioperl (v1.6.1)<sup>41</sup> and PopGenome (v2.0.8)<sup>42</sup>. Consecutive sliding windows with fixed sizes along the genome were adopted for the calculation. Correlation coefficients between two SNPs were calculated by HaploView (v4.2)<sup>43</sup>. A random set of 500,000 SNPs from each breed was used for genome-wide LD estimation.

**Prior genes and ortholog mapping.** Prior genes involved in cardiovascular diseases were retrieved from QIAGEN's Ingenuity Pathway Knowledge Base (accessed 2014-09-28), which is a data repository of manually reviewed biological interactions and functional annotations. In addition, cardiovascular disease-associated genes provided by the NHGRI GWAS Catalog<sup>44</sup> were also incorporated. The human genes were mapped to matched rabbit genes using orthologous information from TreeFam (v9)<sup>45</sup> and Ensembl (v76)<sup>19</sup>.

**RNA expression quantification.** Quality of RNA sequencing reads were checked by the NGS QC Toolkit (v2.3.2)<sup>46</sup>. Reads were filtered if the quality score of more than 75% bases was below 20. Bases from the 3' end were trimmed if their quality was below 20, and reads shorter than 40 bases were discarded after trimming. High quality reads were mapped to the rabbit reference genome (OryCun2) by TopHat2 (v2.0.8)<sup>47</sup>, using the Ensembl transcript annotations (v76)<sup>19</sup> as the reference. Transcripts were assembled and merged by Cufflinks (v2.0.2)<sup>48</sup> with the guidance of the reference. Seven samples with extremely large gene numbers (>50,000) were ignored in the assembly merging. Assembled genes and transcripts with extremely low expression level (FPKM <0.5) in all samples were removed. Reads that mapped to genes were counted by HTSeq (v0.6.0)<sup>49</sup>. Raw counts were normalized by DESeq (v1.16.0)<sup>50</sup>, and DEGs based on the negative binomial distribution were detected with this package (FDR <0.1). The magnitude of expression changes was measured by the log<sub>2</sub>-fold change of normalized counts. Hierarchical clustering was performed with the heatmap.2 function in R (<http://www.r-project.org/>).

**Functional analysis of DEGs.** Function, pathway and regulator enrichment of DEGs were analyzed through the use of QIAGEN's Ingenuity Pathway Analysis (IPA, <http://www.ingenuity.com/products/ipa>, accessed 2014-11-26). Human annotations of the rabbit orthologs were used for analysis of the data sets. IPA's default cut-offs were adopted for activation state prediction (activated: activation z-score >2; inhibited: activation z-score <-2; overrepresented:  $P$ -value < 0.05). The regulation network of lipid metabolism in the liver was visualized by Cytoscape (v3.2.0)<sup>51</sup> and its plugin enhancedGraphics (v1.0.1)<sup>52</sup>.

## References

- Suckow, M. A., Stevens, K. A. & Wilson, R. P. *The laboratory rabbit, guinea pig, hamster, and other rodents*, 157–552 (Academic Press/Elsevier, London; Waltham, MA, 2012).
- Fan, J. *et al.* Rabbit models for the study of human atherosclerosis: From pathophysiological mechanisms to translational medicine. *Pharmacol Ther* **146C**, 104–119 (2015).
- Ignatowski, A. C. Influence of animal food on the organism of rabbits. *S. Peterb. Izviest. Imp. Voenno-Med. Akad.* **16**, 154–173 (1908).
- Watanabe, Y. Serial inbreeding of rabbits with hereditary hyperlipidemia (WHHL-rabbit). *Atherosclerosis* **36**, 261–8 (1980).
- Yamamoto, T., Bishop, R. W., Brown, M. S., Goldstein, J. L. & Russell, D. W. Deletion in cysteine-rich region of LDL receptor impedes transport to cell surface in WHHL rabbit. *Science* **232**, 1230–7 (1986).
- Goldstein, J. L., Kita, T. & Brown, M. S. Defective lipoprotein receptors and atherosclerosis. Lessons from an animal counterpart of familial hypercholesterolemia. *N Engl J Med* **309**, 288–96 (1983).
- Endo, A. Regulation of cholesterol synthesis, as focused on the regulation of HMG-CoA reductase *Seikagaku* **52**, 1033–49 (1980).
- Endo, A. A gift from nature: the birth of the statins. *Nat Med* **14**, 1050–2 (2008).
- Shiomi, M., Koike, T. & Ito, T. Contribution of the WHHL rabbit, an animal model of familial hypercholesterolemia, to elucidation of the anti-atherosclerotic effects of statins. *Atherosclerosis* **231**, 39–47 (2013).
- Carneiro, M. *et al.* The genomic architecture of population divergence between subspecies of the European rabbit. *PLoS Genet* **10**, e1003519 (2014).
- Carneiro, M. *et al.* Rabbit genome analysis reveals a polygenic basis for phenotypic change during domestication. *Science* **345**, 1074–9 (2014).
- Hein, J., Schierup, M. & Wiuf, C. *Gene genealogies, variation and evolution: a primer in coalescent theory*, (Oxford university press, 2004).
- Kumar, P., Henikoff, S. & Ng, P. C. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nat Protoc* **4**, 1073–81 (2009).
- Stachowicz, A. *et al.* Mitochondrial aldehyde dehydrogenase activation by Alda-1 inhibits atherosclerosis and attenuates hepatic steatosis in apolipoprotein E-knockout mice. *J Am Heart Assoc* **3**, e001329 (2014).
- Takeuchi, F. *et al.* Genome-wide association study of coronary artery disease in the Japanese. *Eur J Hum Genet* **20**, 333–40 (2012).
- Corrales, I., Ramirez, L., Altisent, C., Parra, R. & Vidal, F. Rapid molecular diagnosis of von Willebrand disease by direct sequencing. Detection of 12 novel putative mutations in VWF gene. *Thrombosis and Haemostasis* **101**, 570–576 (2009).
- Mori, Y. *et al.* Hypercoagulable State in the Watanabe Heritable Hyperlipidemic Rabbit, an Animal-Model for the Progression of Atherosclerosis - Effect of Probuco on Coagulation. *Thrombosis and Haemostasis* **61**, 140–143 (1989).
- Chen, M. Y. *et al.* Increased expression of lectinlike oxidized low density lipoprotein receptor-1 in initial atherosclerotic lesions of Watanabe heritable hyperlipidemic rabbits. *Arteriosclerosis Thrombosis and Vascular Biology* **20**, 1107–1115 (2000).
- Flicek, P. *et al.* Ensembl 2014. *Nucleic Acids Res* **42**, D749–55 (2014).
- Moore, K. J. & Tabas, I. Macrophages in the pathogenesis of atherosclerosis. *Cell* **145**, 341–55 (2011).
- Desvergne, B., Michalik, L. & Wahli, W. Transcriptional regulation of metabolism. *Physiol Rev* **86**, 465–514 (2006).



22. Hong, C. & Tontonoz, P. Liver X receptors in lipid metabolism: opportunities for drug discovery. *Nat Rev Drug Discov* **13**, 433–44 (2014).
23. Horton, J. D., Goldstein, J. L. & Brown, M. S. SREBPs: activators of the complete program of cholesterol and fatty acid synthesis in the liver. *J Clin Invest* **109**, 1125–31 (2002).
24. Ferre, P. The biology of peroxisome proliferator-activated receptors: relationship with lipid metabolism and insulin sensitivity. *Diabetes* **53** Suppl 1, S43–50 (2004).
25. Chen, C. H., Sun, L. & Mochly-Rosen, D. Mitochondrial aldehyde dehydrogenase and cardiac diseases. *Cardiovasc Res* **88**, 51–7 (2010).
26. Chen, C. H. *et al.* Activation of aldehyde dehydrogenase-2 reduces ischemic damage to the heart. *Science* **321**, 1493–5 (2008).
27. Shiomi, M., Ito, T., Yamada, S., Kawashima, S. & Fan, J. Development of an animal model for spontaneous myocardial infarction (WHHLMI rabbits). *Arterioscler Thromb Vasc Biol* **23**, 1239–44 (2003).
28. Zhang, B. *et al.* Insulin resistance observed in WHHL rabbits. *Atherosclerosis* **91**, 277–8 (1991).
29. Shiomi, M., Kobayashi, T., Kuniyoshi, N., Yamada, S. & Ito, T. Myocardial infarction-prone Watanabe heritable hyperlipidemic rabbits with mesenteric fat accumulation are a novel animal model for metabolic syndrome. *Pathobiology* **79**, 329–38 (2012).
30. Ross, R. Atherosclerosis—an inflammatory disease. *N Engl J Med* **340**, 115–26 (1999).
31. Yu, Y. *et al.* Temporal and quantitative analysis of expression of metalloproteinases (MMPs) and their endogenous inhibitors in atherosclerotic lesions. *Histol Histopathol* **23**, 1503–16 (2008).
32. Kainuma, M. *et al.* Cholesterol-fed rabbit as a unique model of nonalcoholic, nonobese, non-insulin-resistant fatty liver disease with characteristic fibrosis. *Journal of Gastroenterology* **41**, 971–980 (2006).
33. Li, H. & Durbin, R. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* **26**, 589–95 (2010).
34. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–9 (2009).
35. Albers, C. A. *et al.* Dindel: accurate indel calls from short-read data. *Genome Res* **21**, 961–73 (2011).
36. Danecek, P. *et al.* The variant call format and VCFtools. *Bioinformatics* **27**, 2156–8 (2011).
37. Wang, K., Li, M. & Hakonarson, H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res* **38**, e164 (2010).
38. Manichaikul, A. *et al.* Robust relationship inference in genome-wide association studies. *Bioinformatics* **26**, 2867–73 (2010).
39. Yang, J., Lee, S. H., Goddard, M. E. & Visscher, P. M. GCTA: a tool for genome-wide complex trait analysis. *Am J Hum Genet* **88**, 76–82 (2011).
40. Lee, T. H., Guo, H., Wang, X., Kim, C. & Paterson, A. H. SNPhylo: a pipeline to construct a phylogenetic tree from huge SNP data. *BMC Genomics* **15**, 162 (2014).
41. Stajich, J. E. *et al.* The Bioperl toolkit: Perl modules for the life sciences. *Genome Res* **12**, 1611–8 (2002).
42. Pfeifer, B., Wittelsburger, U., Ramos-Onsins, S. E. & Lercher, M. J. PopGenome: an efficient Swiss army knife for population genomic analyses in R. *Mol Biol Evol* **31**, 1929–36 (2014).
43. Barrett, J. C., Fry, B., Maller, J. & Daly, M. J. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* **21**, 263–5 (2005).
44. Welter, D. *et al.* The NHGRI GWAS Catalog, a curated resource of SNP-trait associations. *Nucleic Acids Res* **42**, D1001–6 (2014).
45. Schreiber, F., Patricio, M., Muffato, M., Pignatelli, M. & Bateman, A. TreeFam v9: a new website, more species and orthology-on-the-fly. *Nucleic Acids Res* **42**, D922–5 (2014).
46. Patel, R. K. & Jain, M. NGS QC Toolkit: a toolkit for quality control of next generation sequencing data. *Plos One* **7**, e30619 (2012).
47. Kim, D. *et al.* TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol* **14**, R36 (2013).
48. Trapnell, C. *et al.* Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat Biotechnol* **28**, 511–5 (2010).
49. Anders, S., Pyl, P. T. & Huber, W. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**, 166–9 (2015).
50. Anders, S. & Huber, W. Differential expression analysis for sequence count data. *Genome Biol* **11**, R106 (2010).
51. Cline, M. S. *et al.* Integration of biological networks and gene expression data using Cytoscape. *Nature Protocols* **2**, 2366–2382 (2007).
52. Morris, J. H., Kuchinsky, A., Ferrin, T. E. & Pico, A. R. enhancedGraphics: a Cytoscape app for enhanced node graphics. *F1000Res* **3**, 147 (2014).

## Acknowledgements

We thank Jared Roach and Gustavo Glusman from Institute for Systems Biology for valuable suggestions. This work was supported by the National Basic Research Program of China (2011CB910204, 2011CB510102, 2010CB529200), the Strategic Priority Research Program of the Chinese Academy of Sciences (XDB13040700, XDA12000000), the National Natural Science Foundation of China (NSFC) (91539124, 91439103, 31301032), the National Key Technology Support Program (2013BAI101B09), the National Key Scientific Instrument and Equipment Development Project (2012YQ03026108), the SIBS Knowledge Innovation Program (2014KIP215), the National Institutes of Health grants (HL129778 and HL117491 to Y.E.C.) as well as the Grants-in-Aid for Scientific Research from the Ministry of Education, Culture, Sports and Technology, Japan (22390068, 25670190 and 15H04718 to J.F., and 23300157 to M.S.).

## Author Contributions

Y.L., Y.E.C., J.F., E.L. and L.C. designed the study. J.Z., M.N., J.X., D.Y., J.S., T.Z., S.L., B.N., Z.W., T.K. and M.S. collected samples. G.D. and Z.X. performed the sequencing experiments and analyses. Z.W., Y.D., T.G., S.L., L.Z., L.Z., X.L., J.W., J.Y. and Y.F. performed the population genomics and comparative genomics analyses. Z.W., H.L., J.L., S.W. and M.C. performed the RNA-Seq analyses. J.Z., J.X., D.Y., J.S. and T.Z. performed the validation experiments. H.C. and H.Z. performed the protein structure analyses. Z.W., J.F., J.Z. and H.L. wrote the manuscript.

## Additional Information

**Accession codes:** All sequencing data from this study have been submitted to the NCBI Sequence Read Archive (SRA; <http://www.ncbi.nlm.nih.gov/sra>) under accession number SRP053211 (genome sequencing) and SRP053164 (transcriptome sequencing).

**Supplementary information** accompanies this paper at <http://www.nature.com/srep>

**Competing financial interests:** The authors declare no competing financial interests.

**How to cite this article:** Wang, Z. *et al.* Hyperlipidemia-associated gene variations and expression patterns revealed by whole-genome and transcriptome sequencing of rabbit models. *Sci. Rep.* **6**, 26942; doi: 10.1038/srep26942 (2016).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>