


RESEARCH ARTICLE

Open Access



# Determination of genetic effects and functional SNPs of bovine *HTR1B* gene on milk fatty acid traits

Mingyue Cao<sup>1†</sup>, Lijun Shi<sup>1,2†</sup>, Peng Peng<sup>1</sup>, Bo Han<sup>1</sup>, Lin Liu<sup>3</sup>, Xiaoqing Lv<sup>3</sup>, Zhu Ma<sup>3</sup>, Shengli Zhang<sup>1\*</sup>  and Dongxiao Sun<sup>1\*</sup>

## Abstract

**Background:** Our previous genome-wide association study (GWAS) on milk fatty acid traits in Chinese Holstein cows revealed, the SNP, BTB-01556197, was significantly associated with C10:0 at genome-wide level ( $P = 0.0239$ ). It was located in the down-stream of 5-hydroxytryptamine receptor 1B (*HTR1B*) gene that has been shown to play an important role in the regulation of fatty acid oxidation. Hence, we considered it as a promising candidate gene for milk fatty acids in dairy cattle. In this study, we aimed to investigate whether the *HTR1B* gene had significant genetic effects on milk fatty acid traits.

**Results:** We re-sequenced the entire coding region and 3000 bp of 5' and 3' flanking regions of *HTR1B* gene. A total of 13 SNPs was identified, containing one in 5' flanking region, two in 5' untranslated region (UTR), two in exon 1, five in 3' UTR, and three in 3' flanking region. By performing genotype-phenotype association analysis with SAS9.2 software, we observed that 13 SNPs were significantly associated with medium-chain saturated fatty acids such as C6:0, C8:0 and C10:0 ( $P < 0.0001 \sim 0.042$ ). With Haploview 4.1 software, linkage disequilibrium (LD) analysis was performed. Two haplotype blocks formed by two and ten SNPs were observed. Haplotype-based association analysis indicated that both haplotype blocks were strongly associated with C6:0, C8:0 and C10:0 as well ( $P < 0.0001 \sim 0.0071$ ). With regards to the missense mutation in exon 1 (g.17303383G > T) that reduced amino acid change from alanine to serine, we predicted that it altered the secondary structure of *HTR1B* protein with SOPMA. In addition, we predicted that three SNPs in promoter region, g.17307103A > T, g.17305206 T > G and g.17303761C > T, altered the binding sites of transcription factors (TFs) HMX2, PAX2, FOXP1ES, MIZ1, CUX2, DREAM, and PPAR-RXR by Genomatix. Of them, luciferase assay experiment further confirmed that the allele T of g.17307103A > T significantly increased the transcriptional activity of *HTR1B* gene than allele A ( $P = 0.0007$ ).

**Conclusions:** In conclusion, our findings provided first evidence that the *HTR1B* gene had significant genetic effects on milk fatty acids in dairy cattle.

**Keywords:** Genetic effects, *HTR1B*, Milk fatty acids, Dairy cattle, Luciferase assay

\* Correspondence: [zhangslcau@cau.edu.cn](mailto:zhangslcau@cau.edu.cn); [sundx@cau.edu.cn](mailto:sundx@cau.edu.cn)

<sup>†</sup>Mingyue Cao and Lijun Shi contributed equally to this work.

<sup>1</sup>Department of Animal Genetics, Breeding and Reproduction, College of Animal Science and Technology, Key Laboratory of Animal Genetics, Breeding and Reproduction of Ministry of Agriculture and Rural Affairs, National Engineering Laboratory for Animal Breeding, China Agricultural University, Beijing 100193, China

Full list of author information is available at the end of the article



© The Author(s). 2021 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

## Background

Milk fat, a vital nutritional ingredient of milk, is considered as one of the economic traits of milk production in dairy cattle [1]. Triglyceride synthesized by fatty acids and  $\alpha$ -glycerophosphate in mammary epithelial cells is the main component of milk fat [2]. Fatty acids contain saturated fatty acids (SFAs) and unsaturated fatty acids (UFAs). In SFAs, C12:0, C14:0, C16:0 increases low-density lipoprotein cholesterol and risk of cardiovascular diseases [3], while C15:0, C17:0 are inversely associated with cardiometabolic risk [4]. UFAs are beneficial for reducing the risk of heart and other diseases [5, 6]. Many previous studies have shown that the phenotypic variation of milk fatty acid compositions were genetically controlled and the heritability estimates were around 0.14 ~ 0.33 for SFAs and 0.08 ~ 0.29 for UFAs in Holstein cattle [7–11].

One of genome-wide significant SNPs, BTB-01556197 associated with C10:0 ( $P = 0.0239$ ) identified in previous GWAS [12], was located in the down-stream of 5-hydroxytryptamine receptor 1B (*HTR1B*) gene. The bovine *HTR1B* gene has merely one exon spanning 4305 bp and was involved in the c-AMP signaling pathway that was related to PI3K-Akt pathway, a well-known pathway for fat synthesis and metabolism [13]. Hence, we considered the *HTR1B* gene as a promising candidate for milk fatty acid traits in dairy cattle. In the present study, we aimed to further preform association analysis in a different Chinese Holstein population to confirm the genetic effects of *HTR1B* on milk fatty acids and to identify potential functional genetic variations.

## Results

### SNPs identification

A total of 13 SNPs (Table 1) was identified, including one (g.17307103A > T) in 5' flanking region, two (g.17305206 T > G and g.17303761C > T) in 5' untranslated region (UTR), two (g.17303383G > T and g.17303042C > G) in the exon 1, five (g.17302291C > G, g.17302078G > T, g.17301689 T > C, g.17301647G > T and g.17299803A > G) in 3' UTR, and three (g.17298882A > G, g.17296725C > T and g.17296695 T > C) in 3' flanking region. Of note, g.17303383G > T in exon 1 was a missense mutation resulting in an amino acid replacement from alanine (GCC) to serine (TCC). The detailed information, and genotypic and allele frequencies of the 13 SNPs were shown in Table 1.

### Associations between the SNPs with 23 milk fatty acid traits

By performing association analysis on the SNPs with 23 kinds of milk fatty acids with SAS9.2, significant genetic associations were observed (Additional file 1: Table S1). The SNP, g.17307103A > T, was significantly associated

with C6:0, C8:0, C10:0, C16:1, C16 index and C17 index ( $P = 0.0003 \sim 0.042$ ). Both g.17305206 T > G and g.17303761C > T were significantly associated with C6:0, C10:0, C12:0, C14:0, C18:1cis-9, C18 index, SFA, UFA and SFA/UFA ( $P < 0.0001 \sim 0.036$ ), and they were significantly associated with three (C8:0, C16:1 and C20:0;  $P < 0.0001 \sim 0.0213$ ) and two (C17:0,  $P = 0.0358$ ; and C17 index,  $P = 0.0148$ ) milk fatty acid traits, respectively. Six SNPs were significantly associated with seven milk fatty acid traits, namely, g.17303383G > T (C6:0, C8:0, C10:0, C14:0, C16:1, C16 index, and C17 index;  $P = 0.0003 \sim 0.0145$ ), g.17303042C > G (C8:0, C10:0, C14:0, C16:1, C17:0, C16 index, and C17 index;  $P < 0.0001 \sim 0.0286$ ), g.17302291C > G (C6:0, C8:0, C10:0, C17:0, SFA and SFA/UFA;  $P = 0.0172 \sim 0.0387$ ), g.17302078G > T (C6:0, C8:0, C10:0, C14:0, C16:1, C16 index and C17 index;  $P < 0.0001 \sim 0.0277$ ), g.17299803A > G (C6:0, C8:0, C10:0, C14:0, C16:1, C16 index and C17 index;  $P = 0.001 \sim 0.0305$ ), and g.17296725C > T (C6:0, C8:0, C10:0, C14:0, C16:1, C16 index and C17 index;  $P < 0.0001 \sim 0.0122$ ). The g.17301689 T > C showed strong associations with C6:0, C8:0, C10:0, 16:1, C16 index and C17 index ( $P = 0.0002 \sim 0.0204$ ). The g.17301647G > T, showed strong associations with C6:0, C10:0, C12:0, C14:0, C17:0, C18:1cis-9, C17 index, SFA, UFA and SFA/UFA ( $P < 0.0001 \sim 0.0445$ ). The g.17298882A > G was significantly associated with C14:0, C17:0, C17:1, C18:1cis-9, C20:0, SFA, UFA and SFA/UFA ( $P < 0.0001 \sim 0.049$ ). The g.17296695 T > C was significantly associated with C6:0, C8:0, C10:0, C14:0, C16:1, C17:0, C16 index and C17 index ( $P < 0.0001 \sim 0.0492$ ). While no significant association was observed for C11:0, C13:0, C14:1, C15:0, C16:0, C18:0 and C14 index ( $P > 0.05$ ). After multiple-testing, eight SNPs were still significantly associated with five milk fatty acid traits, namely, g.17305206 T > G (C6:0, C8:0, C10:0, UFA;  $P < 0.0001$ ), g.17303761C > T (C6:0;  $P < 0.0001$ ), g.17303042C > G (C8:0;  $P < 0.0001$ ), g.17302078G > T (C8:0;  $P < 0.0001$ ), g.17301647G > T (C6:0;  $P < 0.0001$ ), g.17298882A > G (C20:0;  $P < 0.0001$ ), g.17296725C > T (C8:0;  $P < 0.0001$ ) and g.17296695 T > C (C8:0;  $P < 0.0001$ ).

Correspondingly, the additive ( $a$ ), dominance ( $d$ ), and allele substitution ( $\alpha$ ) effects of the identified SNPs were calculated (Additional file 2: Table S2), and the results showed that 11 SNPs had strong genetic effects on C6:0, C8:0, C10:0, C12:0, C14:0, C16:1, C17:0, C18:0, C18:1cis-9, C18 index, C20:0, C16 index, C17 index, SFA, UFA and SFA/UFA ( $P < 0.05$ ).

### Associations between the haplotype blocks with 23 milk fatty acid traits

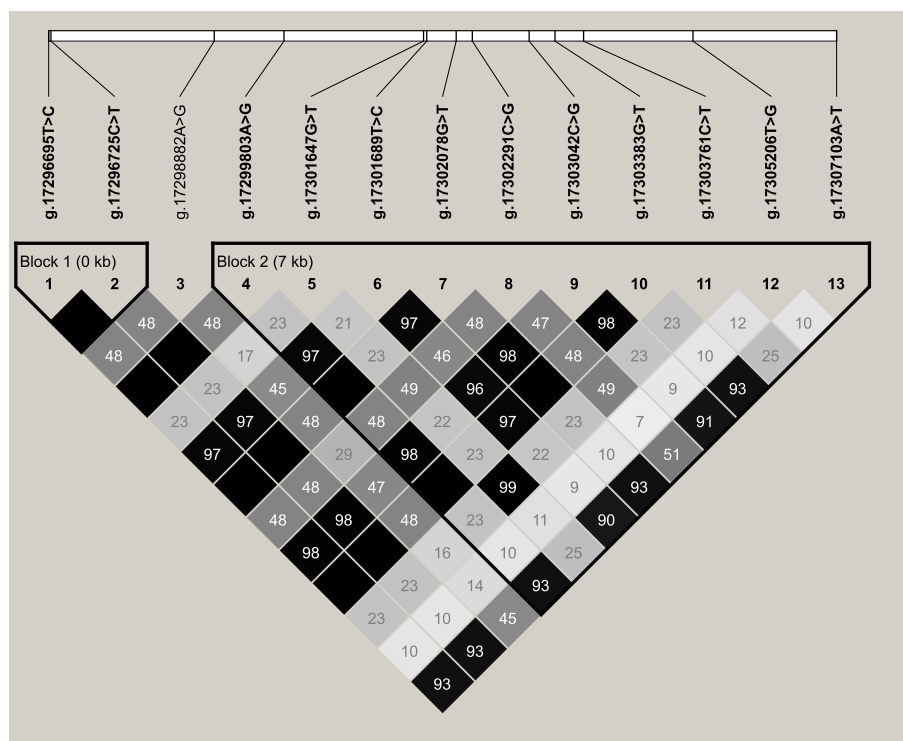
Among the 13 SNPs identified in this study, two haplotype blocks were observed with the Haploview 4.1 software (Fig. 1). The block 1, formed by g.17296725C > T

**Table 1** Detailed information of 13 SNPs identified in *HTR1B* gene

SNP name	Location	Position (UMD 3.1.1)	EVA no.	Genotype	NO.	Frequency	Allele	Frequency
g.17307103A > T	5' flanking region	Chr9: 17307103	rs207969357	AA	481	0.4621	A	0.6830
				TT	100	0.0961	T	0.3170
				TA	460	0.4419		
g.17305206 T > G	5' UTR	Chr9: 17305206	rs476055046	TT	651	0.6242	T	0.8121
				GG	0	0.0000	G	0.1879
				GT	392	0.3758		
g.17303761C > T	5' UTR	Chr9: 17303761	rs133683693	CC	113	0.1088	C	0.3503
				TT	424	0.4081	T	0.6497
				CT	502	0.4832		
g.17303383G > T	Exon-1	Chr9: 17303383	rs209984404	GG	504	0.4846	G	0.6986
				TT	91	0.0875	T	0.3014
				GT	445	0.4279		
g.17303042C > G	Exon-1	Chr9: 17303042	rs208945882	CC	499	0.4897	C	0.6977
				GG	96	0.0942	G	0.3023
				CG	424	0.4161		
g.17302291C > G	3' UTR	Chr9: 17302291	rs136136524	CC	279	0.2680	C	0.5259
				GG	225	0.2161	G	0.4741
				CG	537	0.5159		
g.17302078G > T	3' UTR	Chr9: 17302078	rs135063494	GG	502	0.4818	G	0.6972
				TT	91	0.0873	T	0.3028
				GT	449	0.4309		
g.17301689 T > C	3' UTR	Chr9: 17301689	rs379078023	CC	103	0.1012	C	0.3089
				TT	492	0.4833	T	0.6911
				CT	423	0.4155		
g.17301647G > T	3' UTR	Chr9: 17301647	rs109548495	GG	115	0.1103	G	0.3514
				TT	425	0.4075	T	0.6486
				GT	503	0.4823		
g.17299803A > G	3' UTR	Chr9: 17299803	rs208790360	AA	499	0.4854	A	0.6989
				GG	90	0.0875	G	0.3011
				AG	439	0.4270		
g.17298882A > G	3' flanking region	Chr9: 17298882	rs380460000	AA	2	0.0038	A	0.5019
				GG	0	0.0000	G	0.4981
				AG	523	0.9962		
g.17296725C > T	3' flanking region	Chr9: 17296725	rs208087947	CC	503	0.4846	C	0.6985
				TT	91	0.0877	T	0.3015
				CT	444	0.4277		
g.17296695 T > C	3' flanking region	Chr9: 17296695	rs133617481	CC	90	0.0872	C	0.2994
				TT	504	0.4884	T	0.7006
				CT	438	0.4244		

Note: EVA European Variation Archive, UTR untranslated region

and g.17296695 T > C, had two kinds of haplotypes (H1 = TC and H2 = CT) with the frequencies of 69.6 and 30.4%, respectively. The block 2 was composed of ten SNPs, g.17307103A > T, g.17305206 T > G, g.17303761C > T, g.17303383G > T, g.17303042C > G, g.17302291C > G, g.17302078G > T, g.17301689 T > C, g.17301647G > T and g.17299803A > G, and had seven haplotypes (H1 = AGTGCCGCTA, 34.5%; H2 =



**Fig. 1** Linkage disequilibrium (LD) among the 13 SNPs of *HTR1B* gene

GTCTGGTTTT, 30.3%; H3 = ATTGGCGTGA, 13.8%; H4 = ATTGCCGTGA, 12.9%; H5 = ATTGCCGTGA, 4.4%; H6 = ATTGGCGTGA, 1.6%; and H7 = ATTGGCGTTT, 1.5%).

Subsequently, we performed haplotype-based association analysis with SAS9.2, and found that the haplotype block1 and block2 were significantly associated with nine (C6:0, C8:0, C10:0, C14:0, C16:1, 17:0, C20:0, C16 index and C17 index;  $P = 0.0002 \sim 0.0476$ ) and five (C6:0, C8:0, C10:0, C20:0 and C17 index;  $P < 0.0001 \sim 0.0265$ ) milk fatty acids, respectively (Additional file 3: Table S3). While, none of significant associations were detected with C11:0, C12:0, C13:0, C14:1, C15:0, C16:0, 17:1, C18:0, C18:1cis-9, C18 index, C14 index, SFA, UFA and SFA/UFA ( $P > 0.05$ ). Through multiple testing, the haplotype block1 and block2 were still significantly associated with one (C8:0;  $P = 0.0002$ ) and three (C6:0, C8:0, C10:0;  $P < 0.0001$ ) milk fatty acids, respectively.

**Changes of the HTR1B protein secondary structure and function caused by the missense mutation g.17303383G > T**

Using SOPMA software, we predicted that the missense mutation in exon 1, g.17303383G > T, changed the HTR1B protein secondary structure, including  $\alpha$ -helix (46.90 to 41.44%), extended strand (13.40 to 16.38%),  $\beta$ -turn (2.48 to 1.74%), and random coil (37.22 to 40.45%)

with the allele from G to T. While, the HTR1B protein function was not altered by the missense mutation with the scores 0.25 (SIFT) and -1.42 (PROVEAN).

**Changes of transcriptional activity caused by g.17307103A > T, g.17305206T > G and g.17303761C > T**

By searching the TFBSs of the two SNPs in 5' UTR and one SNP in 5' flanking region with Genomatix, we discovered that the allele T of g.17307103A > T created two TFBSs for HMX2 (Hmx2/Nkx5-2 homeodomain transcription factor) and PAX2 (Zebrafish PAX2 paired domain protein), and the allele G of g.17305206T > G created a TFBS for FOXP1ES (Alternative splicing variant of FOXP1, activated in ESCs). For g.17303761C > T, the allele C created two TFBSs for MIZ1 (Myc-interacting Zn finger protein 1, zinc finger and BTB domain containing 17) and CUX2 (Cut-like homeobox 2, dimeric binding site), and the allele T created two TFBSs for DREAM (Downstream regulatory element-antagonist modulator, Ca<sup>2+</sup>-binding protein of the neuronal calcium sensors family that binds DRE sites as a tetramer) and PPAR-RXR (PPAR/RXR heterodimers, DR1 sites). The detailed results were shown in Table 2.

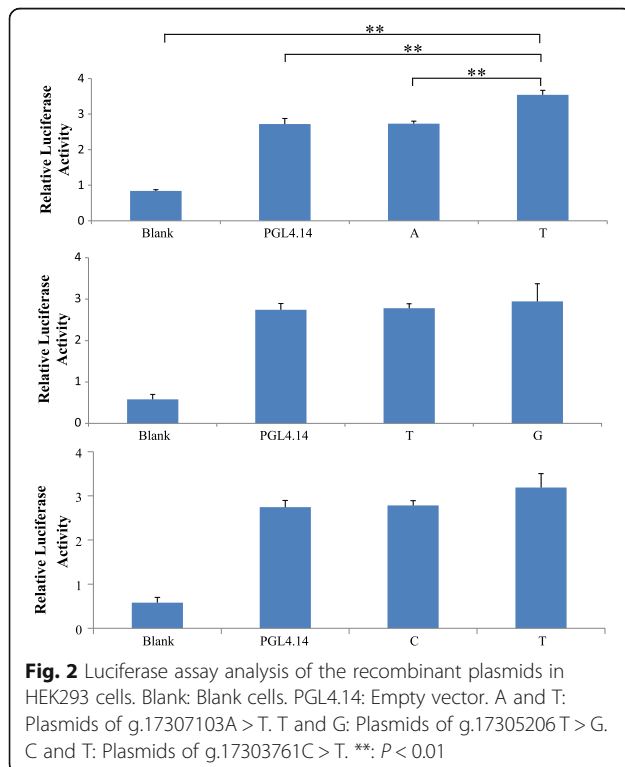
Further, we utilized the luciferase assay (Fig. 2) to confirm the above prediction results. The luciferase activities of the construct T of g.17307103A > T was observed significantly higher than those of the blank

**Table 2** Changes of transcription factor binding sites (TFBSs) caused by the SNPs in the 5' UTR and flanking region of *HTR1B*

SNP	Sequence	TF	Full name
g.17307103A > T	CCCCAACGCGATTC CCTCCTT		
	CCCCAACGCGTTTC CCTCCTT	Hmx2 PAX2	Hmx2/Nkx5-2 homeodomain transcription factor Zebrafish PAX2 paired domain protein
g.17305206 T > G	TTTTGAAGTTTTTTTT TTTTTT		
	TTTTGAAGTTTGTT TTTTTTTT	FOXP1ES	Alternative splicing variant of FOXP1, activated in ESCs
g.17303761C > T	ACCTCGCCCTCGAC CTCTCGC	MIZ1	Myc-interacting Zn finger protein 1, zinc finger and BTB domain containing 17 (ZBTB17)
		CUX2	Cut-like homeobox 2, dimeric binding site
	ACCTCGCCCTTGAC CTCTCGC	DREAM PPAR-RXR	Downstream regulatory element-antagonist modulator, Ca <sup>2+</sup> -binding protein of the neuronal calcium sensors family that binds DRE (downstream regulatory element) sites as a tetramer PPAR/RXR heterodimers, DR1 sites

Notes: TF transcription factor

control ( $P < 0.0001$ ), the empty vector PGL4.14 ( $P = 0.0022$ ), and the construct A ( $P = 0.0007$ ), indicating that the allele T of g.17307103A > T increased the transcriptional activity of *HTR1B* compared with allele A. However, the luciferase activities of four constructs (T and G of g.17305206 T > G, and C and T of g.17303761C > T) were not strongly changed than those of the blank and empty vector, implying that g.17305206 T > G and g.17303761C > T did not significantly alter the transcriptional activity of *HTR1B* gene.



## Discussion

According to our previous GWAS results and biological functions, the *HTR1B* gene has been identified as one of novel promising candidates for milk fatty acids in dairy cattle [12]. In the present study, we firstly confirmed that the *HTR1B* gene showed significant genetic effects on medium-chain saturated fatty acids in dairy cattle, providing a basis for further verification.

Previous studies reported that protein secondary structure could be used to build safe starting cores to generate the complete protein fold [14], and to set structural constraints for protein threading [15, 16]. Studies reported that the missense mutations caused by the sequence variations were related to the protein function to account for the phenotype variations [17–20]. Here, we identified a missense mutation (g.17303383G > T), and it changed the protein secondary structure by prediction with the SOPMA. While we used SIFT and PROVEAN softwares to detect that the missense mutation did not alter the *HTR1B* protein function. Hence, the significant associations of g.17303383G > T with milk fatty acids might be due to the LD with the true causal mutations.

Regulatory region, including promoter, enhancer, silencer, and insulator etc. [21], are important for the gene regulation and expression [22]. Transcription factors (TFs) are the sequence-specific DNA-binding proteins that regulate gene expression in all organisms [23, 24], and approximately 10% are implicated in diverse diseases in human [25]. In eukaryotes, multiple TFs cooperatively bind regulatory DNA to temporally and spatially control gene expression [26]. Generally speaking, 5' UTR plays regulatory roles in the gene expression through binding transcription factors or unknown regulatory mechanisms. Hence, we wish to investigate if the SNPs in the 5' UTRs changed the expression of *HTR1B* gene. It has been well-known that

transcription factors (TFs) play an important role in gene expression and transcriptional regulation [27, 28]. Previous studies found that both 5' UTR and 5' flanking region can combine with TF, and some SNPs within these regions may change the TF-binding thereby leading to changes in gene transcription activity and eventually affecting related traits [29–31]. In this study, we predicted that three SNPs (g.17307103A > T, g.17305206 T > G and g.17303761C > T) changed the TFBSs. While only g.17307103A > T actually changed the transcriptional activity of *HTR1B* gene. The allele T of g.17307103A > T was predicted to create two TFBSs for HMX2 and PAX2. The *hmx* homeobox-containing TF gene family, containing contains *hmx2* [32], is highly conserved across species [33–36]. HMX2 is involved in a feedback loop of EGF signaling and located in upstream of the *PAX5* of the utricular maculae to affect the inner ear development in zebrafish [37, 38]. The *pax* gene family encodes DNA binding TFs that control vital steps in embryonic development and differentiation of specific cell lineages in human [39]. PAX2 as a TF can promote the expression of *ADAM10* to negatively regulate the epithelia to mesenchyme transition in human [40]. According to the significant associations and transcriptional activity caused by g.17307103A > T, we suggested that g.17307103A > T might be a potential causal mutation regulating the *HTR1B* gene expression by altering the binding sites for the TFs HMX2 and PAX2 to impact the milk fatty acid traits in dairy cattle. Further investigation is needed to validate the regulatory of specific transcription factors.

Regarding the expression of the *HTR1B* gene in multiple tissues, based on the RNA-seq database, Cattle Gene Atlas (<http://cattlegeneatlas.roslin.ed.ac.uk/>) [41], we observed that the *HTR1B* gene was expressed in 82 tissues/ cell types, including mammary gland, while *HTR1B* gene was moderately expressed in mammary gland.

## Conclusions

In this study, we confirmed the significant genetic effects of the *HTR1B* gene on milk fatty acids using post-GWAS strategy, and identified a potential functional mutation in 5'-flanking region, g.17307103A > T, that altered the transcriptional activity of *HTR1B*. Our findings provided valuable molecular information for genetic improvement programs in dairy cattle.

## Methods

### Animals and phenotypic data

A total of 1065 Chinese Holstein cows was used in this study that were different from those in our previous GWAS [12]. They were from 44 sire families with an average of 24 daughters per sire and collected in the 23 dairy farms of the Beijing Sanyuanlyhe Dairy Farming Center (Beijing, China). We collected milk samples of these 1065 cows to measure milk fatty acids in the laboratory of Beijing Dairy Cattle

Center ([www.bdcc.com.cn](http://www.bdcc.com.cn)). By gas chromatograph, the contents of 16 kinds of main milk fatty acids were directly detected, including SFA: C6:0, C8:0, C10:0, C11:0, C12:0, C13:0, C14:0, C15:0, C16:0, C17:0, C18:0, C20:0; and UFA: C14:1, C16:1, C17:1 and C18:1cis-9. Based on these phenotype values, we obtained four index traits (C14 index, C16 index, C17 index and C18 index) calculated with the formulas:  $\frac{\text{cis-9 unsaturated}}{\text{cis-9 unsaturated} + \text{saturated}} * 100$  [42], and summarized the SFA, UFA, SFA/UFA.

### Gas chromatograph

Total milk fat were extracted from approximately 2 ml of each milk sample. Fatty acids have high boiling points, so they are unstable and easy to crack at high temperatures. The high temperature in gas chromatographic analysis will cause the loss of fatty acids, so pre-treatment is required. Therefore, when analyzing fatty acids and fats, especially fatty acid components, to reduce the boiling point and improve stability, we react the fatty acids or fats with methanol to prepare fatty acid methyl esters and then perform gas chromatography analysis.

The measurement conditions of the gas chromatograph are as follows: the injector temperature is 260 °C; carrier gas (helium) flow rate is 45 mL/min; split ratio is 100:1; chromatographic column conditions: keep at 100 °C for 10 min, heat up at 6 °C /min to 160 °C and hold for 10 min, heat up at 5 °C /min to 200 °C and hold for 20 min, heat up at 4 °C /min to 240 °C and hold for 12 min; the detector temperature is 260 °C.

### SNP identification and genotyping

We extracted genomic DNA from the blood samples of 1065 cows and the semen samples of 44 sires using TIA-Namp Blood DNA Kit (Tiangen, Beijing, China) and salt-out procedure, respectively. DNA quantity and quality were measured by NanoDrop™ ND-2000 Spectrophotometer (Thermo Scientific, Hudson, DE, USA) and 2.0% agarose gel electrophoresis.

A total of 15 pairs of primers (Additional file 4: Table S4) were designed for PCR amplification using the Primer 3.0 (<http://primer3.wi.mit.edu/>) based on the sequences of all the exons, and 3000 bp of 5' and 3' flanking regions of the bovine *HTR1B* gene (Gene ID: 317707), and were synthesized in the Beijing Genomics Institute (Beijing, China). By using the DNA samples of the abovementioned 44 sires with the same concentration of 50 ng/μl, two DNA pools were constructed and 22 sires were included in each pool. Then, PCR amplification was performed with abovementioned 15 pairs of primers and PCR procedure was as follows: initial denaturation at 94 °C for 5 min; annealing at 94 °C for 30 s, 60 °C for 30 s and 72 °C for 30 s, for 35 cycles and final extension at 72 °C for 7 min. By sequencing PCR products, we identified potential polymorphic sites.

Then, 1065 cows were individually genotyped by using matrix-assisted laser desorption/ionization time of flight mass spectrometry (MALDI-TOF MS, Sequenom MassARRAY, Bioyong Technologies Inc. HK). As for each identified SNP, PCR amplification was first performed with sequence-specific extension primers, then 1 base was extended targeting two alleles of the identified SNP. According to different mass-to-charge ratios of two alleles, different mass spectrum peaks could be observed to detect the genotype of each SNP.

### Statistical analysis

First, we used the Haploview 4.1 software (Broad Institute of MIT and Harvard, Cambridge, MA, USA) to identify the LD extent among the identified SNPs of the *HTR1B* gene.

Subsequently, we performed single SNP-based and haplotype-based association analysis. We traced the pedigrees of the 1065 Chinese Holstein cows back to three generations, as a result, a total of 3335 individuals were included for association analysis, which kinship matrix (A-matrix) were constructed with SAS9.2 (SAS Institute, Cary, NC, USA). Then, associations between the identified SNPs and haplotype blocks with 23 milk fatty acid traits were performed by SAS9.2 on the basis of the following mixed animal model:

$$Y_{ijklm} = \mu + G_i + h_j + l_k + a_l + b \times M_m + e_{ijklm}$$

Here,  $Y_{ijklm}$  is the phenotypic value of each milk fatty acid trait;  $\mu$  is the overall mean;  $G_i$  is the fixed effect corresponding to the genotype or haplotype combination;  $h_j$  is the fixed effect of farm ( $j = 1 \sim 23$ );  $l_k$  is the fixed effect of stage of lactation ( $k = 1 \sim 4$ );  $a_l$  is the random polygenic effect;  $M_m$  is the fixed effect of age at calving ( $m = 1 \sim 293$ );  $b$  is the regression coefficient of covariate  $M$ ; and  $e_{ijklm}$  is the random residual. Bonferroni correction was performed according to the number of multiple tests, in which the adjusted significance levels of  $P < 0.05$  for the single SNP and haplotype-based analysis were 0.0002 and 0.0011, respectively.

Further, we calculated the additive effect ( $a$ ), dominant effect ( $d$ ), and substitution effect ( $\alpha$ ) of SNP on the milk fatty acid traits according to the formulas [43]:  $a = \frac{AA-BB}{2}$ ,  $d = AB - \frac{AA+BB}{2}$ , and  $\alpha = a + d(q-p)$ , in which,  $p$  and  $q$  were the frequencies of A and B, respectively; and AA, AB and BB were the least square means of fatty acids corresponding to the genotypes.

### Prediction of the secondary structure and function changes of the HTR1B protein

We used the NPSA SOPMA SERVER program ([https://npsa-prabi.ibcp.fr/cgi-bin/npsa\\_automat.pl?page=npsa\\_sopma.html](https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=npsa_sopma.html)) to predict whether the identified missense mutation in coding region changed HTR1B protein secondary structure, and set the parameters with similarity

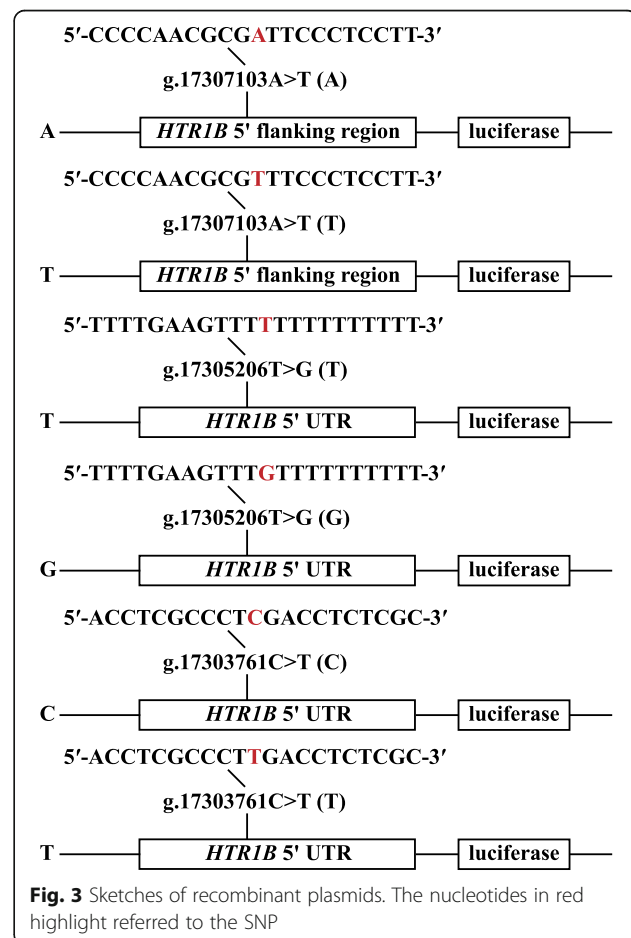
threshold (8), and number of states (4-Helix, Sheet, Turn, Coil). Also, we used the SIFT (<http://sift.bii.a-star.edu.sg/>) and PROVEAN (<http://sift.jcvi.org/index.php>) to investigate whether the missense mutation altered the protein function. The score thresholds of the SIFT and PROVEAN were 0.05 [44] and  $-2.5$  [45], respectively. When the score is below the threshold, the protein function is considered changed.

### Prediction of the changes of transcription factor binding sites (TFBSs)

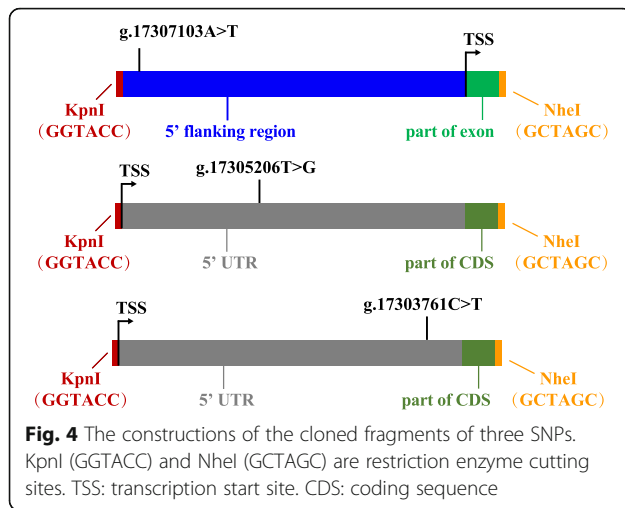
We predicted whether the SNPs in 5' UTR and flanking region of the *HTR1B* gene impacted on TFBSs by using the Genomatix suite v3.9 (<http://www.genomatix.de/cgi-bin/welcome/welcome.pl?s=d1b5c9a9015b02bb3b1a806f9c03293f>).

### Construction of recombinant plasmid, cell culture and luciferase assay

We constructed six luciferase reporter gene fragments with Kpn1 and NheI restriction sites at the 5' to 3' termini (Figs. 3 and 4), which contained alleles A and T of g.17307103A > T, T and G of g.17305206 T > G, and C and



**Fig. 3** Sketches of recombinant plasmids. The nucleotides in red highlight referred to the SNP



T of g.17303761C > T. The six fragments were synthesized in Genewiz (Suzhou, China), and were cloned into the pGL4.14 Luciferase Assay Vector (Promega, Madison, USA). Subsequently, the plasmids were purified by Endo-free Plasmid DNA Mini Kit II (OMEGA, omega bio-tek, Norcross, Georgia, USA), and then were sequenced to confirm the integrity of each construct's insertions.

Human Embryonic Kidney (HEK)-293 T cells were cultured in Dulbecco's modified Eagle's medium (DMEM) including 10% heat-inactivated fetal bovine serum (FBS; Gibco, Life Technologies) at 5% CO<sub>2</sub> and 37 °C. We seeded approximately 2 × 10<sup>5</sup> cells per well into 24-well plates, and transfected the cells using Lipofectamine 2000 (Invitrogen, CA, USA). For each well, we transfected 500 ng of the constructed plasmid DNA along with 10 ng of pRL-TK renilla luciferase reporter vector (Promega), and conducted three replicates for each construct. The cells were cultured for about 36–48 h after transfection and then were measured the activities of firefly and renilla luciferases using a Dual-Luciferase Reporter Assay System (Promega) with a Modulus microplate multimode reader (Turner Biosystems, CA, USA). Finally, average statistics of three replicates were calculated as the normalized luciferase data (firefly/renilla).

#### Abbreviations

a: Additive effect; CDS: coding sequence; d: Dominant effect; EVA: European Variation Archive; FA: Fatty acids; GWAS: Genome-wide association study; *HTR1B*: 5-hydroxytryptamine receptor 1B; LD: Linkage disequilibrium; PCR: Polymerase chain reaction; SFA: Saturated fatty acids; SNP: Single nucleotide polymorphism; TF: Transcription factor; TFBS: Transcription factor binding site; TSS: transcription start site; UFA: Unsaturated fatty acids; UTR: Untranslated region; α: Substitution effect

#### Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12864-021-07893-8>.

**Additional file 1: Table S1.** Associations of 13 SNPs of *HTR1B* gene with fatty acid traits in Chinese Holstein (LSM ± SE).

**Additional file 2: Table S2.** Additive(a), dominant(d) and substitution(α) effects of 11 SNPs on milk fatty acid traits of *HTR1B* gene.

**Additional file 3: Table S3.** Haplotype associations of the haplotype blocks in *HTR1B* gene with milk fatty acids (LSM ± SE).

**Additional file 4: Table S4.** PCR primers information of *HTR1B* gene.

#### Acknowledgements

We appreciate Beijing Dairy Cattle Center, Beijing Municipal Bureau of Agriculture, and Beijing Sanyuanlvhe Dairy Farming Center, for providing the milk, semen, and blood samples of Chinese Holstein.

#### Authors' contributions

DS and SZ conceived and designed the experiments, LL, XL and ZM prepared the milk, blood and semen samples, PP measured the milk fatty acids, MC and LS extracted the DNA for SNP identification and genotyping with the help of BH, MC analyzed the data, and the manuscript was prepared by MC, LS, SZ and DS. All authors read and approved the final manuscript.

#### Funding

This research was funded by the Beijing Science and Technology Program (20200105, D171100002417001), National Science and Technology Programs of China (2013AA102504), National Natural Science Foundation of China (31872330, 31802041), earmarked fund for Modern Agro-industry Technology Research System (CARS-36), and the Program for Changjiang Scholar and Innovation Research Team in University (IRT\_15R62). These funding bodies had no roles in the design of the study and collection, analysis, and interpretation of data and in writing the manuscript.

#### Availability of data and materials

All data generated or analysed during this study are included in this published article and its supplementary information files.

#### Declarations

##### Ethics approval and consent to participate

All protocols for collection of the samples of experimental individuals and phenotypic observations were reviewed and approved by the Institutional Animal Care and Use Committee (IACUC) at China Agricultural University (Permit Number: DK996). Blood, semen and milk samples were collected specifically for this study following standard procedures with the full agreement of the Beijing Sanyuanlvhe Dairy Farming Center who owned the Holstein cows and bulls, respectively.

##### Consent for publication

Not applicable.

##### Competing interests

The authors declare that they have no competing interests.

##### Author details

<sup>1</sup>Department of Animal Genetics, Breeding and Reproduction, College of Animal Science and Technology, Key Laboratory of Animal Genetics, Breeding and Reproduction of Ministry of Agriculture and Rural Affairs, National Engineering Laboratory for Animal Breeding, China Agricultural University, Beijing 100193, China. <sup>2</sup>Institute of Animal Science, Chinese Academy of Agricultural Sciences, Beijing 100193, China. <sup>3</sup>Beijing Dairy Cattle Center, Beijing 100192, China.

Received: 24 June 2021 Accepted: 15 July 2021

Published online: 27 July 2021

#### References

- Spelman RJ, Coppieters W, Karim L, van Arendonk JA, Bovenhuis H. Quantitative trait loci analysis for five milk production traits on chromosome six in the Dutch Holstein-Friesian population. *Genetics*. 1996;144(4):1799–808. <https://doi.org/10.1093/genetics/144.4.1799>.
- Pennington RJ. The metabolism of short-chain fatty acids in the sheep. I. Fatty acid utilization and ketone body production by rumen epithelium and other tissues. *Biochem J*. 1952;51(2):251–8. <https://doi.org/10.1042/bj0510251>.



3. Lordan R, Tsoupras A, Mitra B, Zabetakis I. Dairy Fats and Cardiovascular Disease: Do We Really Need to be Concerned? *Foods*. 2018;7(3):29. <https://doi.org/10.3390/foods7030029>.
4. Yu E, Hu FB. Dairy products, dairy fatty acids, and the prevention of Cardiometabolic disease: a review of recent evidence. *Curr Atheroscler Rep*. 2018;20(5):24. <https://doi.org/10.1007/s11883-018-0724-z>.
5. Assmann KE, Adjibade M, Herberg S, Galan P, Kesse-Guyot E. Unsaturated fatty acid intakes during midlife are positively associated with later cognitive function in older adults with modulating effects of antioxidant supplementation. *J Nutr*. 2018;148(12):1938–45. <https://doi.org/10.1093/jn/nxy206>.
6. Kris-Etherton PM, Pearson TA, Wan Y, Hargrove RL, Moriarty K, Fishell V, et al. High-monounsaturated fatty acid diets lower both plasma cholesterol and triacylglycerol concentrations. *Am J Clin Nutr*. 1999;70(6):1009–15. <https://doi.org/10.1093/ajcn/70.6.1009>.
7. Narayana SG, Schenkel FS, Fleming A, Koeck A, Malchiodi F, Jamrozik J, et al. Genetic analysis of groups of mid-infrared predicted fatty acids in milk. *J Dairy Sci*. 2017;100(6):4731–44. <https://doi.org/10.3168/jds.2016-12244>.
8. Petrini J, lung LH, Rodriguez MA, Salvian M, Pertille F, Rovadoski GA, et al. Genetic parameters for milk fatty acids, milk yield and quality traits of a Holstein cattle population reared under tropical conditions. *J Anim Breed Genet*. 2016;133(5):384–95.
9. Krag K, Poulsen NA, Larsen MK, Larsen LB, Janss LL, Buitenhuis B. Genetic parameters for milk fatty acids in Danish Holstein cattle based on SNP markers using a Bayesian approach. *BMC Genet*. 2013;14(1):79. <https://doi.org/10.1186/1471-2156-14-79>.
10. Stoop WM, Schennink A, Visker MH, Mullaart E, van Arendonk JA, Bovenhuis H. Genome-wide scan for bovine milk-fat composition. I. Quantitative trait loci for short- and medium-chain fatty acids. *J Dairy Sci*. 2009;92(9):4664–75. <https://doi.org/10.3168/jds.2008-1966>.
11. Schennink A, Stoop WM, Visker MH, van der Poel JJ, Bovenhuis H, van Arendonk JA. Short communication: genome-wide scan for bovine milk-fat composition. II. Quantitative trait loci for long-chain fatty acids. *J Dairy Sci*. 2009;92(9):4676–82. <https://doi.org/10.3168/jds.2008-1965>.
12. Li C, Sun DX, Zhang SL, Wang S, Wu XP, Zhang Q, Liu L, Li YH, Qiao L. Genome Wide Association Study Identifies 20 Novel Promising Genes Associated with Milk Fatty Acid Traits in Chinese Holstein. *PLoS One*. 2014; 9(5):e96186. <https://doi.org/10.1371/journal.pone.0096186>.
13. Fan Y, He Z, Wang W, Li J, Hu A, Li L, et al. Tangganjian decoction ameliorates type 2 diabetes mellitus and nonalcoholic fatty liver disease in rats by activating the IRS/PI3K/AKT signaling pathway. *Biomed Pharmacother*. 2018;106:733–7.
14. Colubri A. Prediction of protein structure by simulating coarse-grained folding pathways: a preliminary report. *J Biomol Struct Dyn*. 2004;21(5):625–38. <https://doi.org/10.1080/07391102.2004.10506953>.
15. Rost B, Schneider R, Sander C. Protein fold recognition by prediction-based threading. *J Mol Biol*. 1997;270(3):471–80. <https://doi.org/10.1006/jmbi.1997.1101>.
16. Fischer D, Eisenberg D. Protein fold recognition using sequence-derived predictions. *Protein science: a publication of the Protein Society*. 1996;5(5): 947–55. <https://doi.org/10.1002/pro.5560050516>.
17. Cheng YJ, Yao H, Ji CC, Chen XM, Fan J, Liu LJ, et al. A heterozygous missense hERG mutation associated with early repolarization syndrome. *Cell Physiol Biochem*. 2018;51(3):1301–12. <https://doi.org/10.1159/000495549>.
18. Nakayama M, Oshima M. Mutant p53 in colon cancer. *J Mol Cell Biol*. 2018; 11(4):267–76. <https://doi.org/10.1093/jmcb/mjy075>.
19. Guedes LB, Almutairi F, Haffner MC, Rajoria G, Liu Z, Klimek S, et al. Analytic, Preanalytic, and clinical validation of p53 IHC for detection of TP53 missense mutation in prostate Cancer. *Clin Cancer Res*. 2017;23(16):4693–703. <https://doi.org/10.1158/1078-0432.CCR-17-0257>.
20. Li LX, Dong HL, Xiao BG, Wu ZY. A novel missense mutation in peripheral myelin Protein-22 causes Charcot-Marie-tooth disease. *Chin Med J*. 2017; 130(15):1779–84. <https://doi.org/10.4103/0366-6999.211539>.
21. Singh AP, Mishra S, Jabin S. Sequence based prediction of enhancer regions from DNA random walk. *Sci Rep*. 2018;8(1):15912. <https://doi.org/10.1038/s41598-018-33413-y>.
22. Plank JL, Dean A. Enhancer function: mechanistic and genome-wide insights come together. *Mol Cell*. 2014;55(1):5–14. <https://doi.org/10.1016/j.molcel.2014.06.015>.
23. Lelli KM, Slattery M, Mann RS. Disentangling the many layers of eukaryotic transcriptional regulation. *Annu Rev Genet*. 2012;46(1):43–68. <https://doi.org/10.1146/annurev-genet-110711-155437>.
24. Lambert SA, Jolma A, Campitelli LF, Das PK, Yin Y, Albu M, et al. The human transcription factors. *Cell*. 2018;175(2):598–9. <https://doi.org/10.1016/j.cell.2018.09.045>.
25. Bouhleb MA, Lambert M, David-Cordonnier MH. Targeting transcription factor binding to DNA by competing with DNA binders as an approach for controlling gene expression. *Curr Top Med Chem*. 2015;15(14):1323–58. <https://doi.org/10.2174/1568026615666150413154713>.
26. Chen H, Li H, Liu F, Zheng X, Wang S, Bo X, et al. An integrative analysis of TFBS-clustered regions reveals new transcriptional regulation models on the accessible chromatin landscape. *Sci Rep*. 2015;5(1):8465. <https://doi.org/10.1038/srep08465>.
27. van der Flier LG, van Gijn ME, Hatzis P, Kujala P, Haegebarth A, Stange DE, et al. Transcription factor Achaete Scute-like 2 controls intestinal stem cell fate. *Cell*. 2009;136(5):903–12. <https://doi.org/10.1016/j.cell.2009.01.031>.
28. Francis R, Guo HY, Streutker C, Ahmed M, Yung T, Dirks PB, He HH, Kim TH. Gastrointestinal transcription factors drive lineage-specific developmental programs in organ specification and cancer. *Sci Adv*. 2019;5(12):eaax8898. <https://doi.org/10.1126/sciadv.aax8898>.
29. Wang XY, Li AN, Raza SHA, Liang CC, Zhang S, Mei CG, et al. Transcription factors ZEB1 and CREB promote the transcription of bovine *ABHD5* gene. *DNA Cell Biol*. 2020;40(2):219–230. <https://doi.org/10.1089/dna.2020.5994>.
30. Liu FF, Wang KL, Deng LP, Liu X, Wu MY, Wang TY, Cui L, Li ZG. Transcription factor E2F3a regulates *CASP8A2* transcription and enhances sensitivity to chemotherapeutic drugs in acute lymphoblastic leukemia. *Cancer Cell Int*. 2018;18(1):40. <https://doi.org/10.1186/s12935-018-0531-1>.
31. Zhang X, Li YT, Sun YL, Guo MX, Feng JJ, Wang YL, Zhang ZP. Regulatory effect of heat shock transcription factor-1 gene on heat shock proteins and its transcriptional regulation analysis in small abalone *Haliotis diversicolor*. *Bmc Mol Cell Biol*. 2020;21(1):83. <https://doi.org/10.1186/s12860-020-00323-9>.
32. Adamska M, Wolff A, Kreusler M, Wittbrodt J, Braun T, Bober E. Five Nkx5 genes show differential expression patterns in anlagen of sensory organs in medaka: insight into the evolution of the gene family. *Dev Genes Evol*. 2001;211(7):338–49. <https://doi.org/10.1007/s004270100162>.
33. Stadler HS, Murray JC, Leysens NJ, Goodfellow PJ, Solursh M. Phylogenetic conservation and physical mapping of members of the H6 homeobox gene family. *Mamm Genome*. 1995;6(6):383–8. <https://doi.org/10.1007/BF00355637>.
34. Jiang Y, Horner V, Liu J. The HMX homeodomain protein MLS-2 regulates cleavage orientation, cell proliferation and cell fate specification in the *C. elegans* postembryonic mesoderm. *Development*. 2005;132(18):4119–30. <https://doi.org/10.1242/dev.01967>.
35. Kiernan AE, Nunes F, Wu DK, Fekete DM. The expression domain of two related homeobox genes defines a compartment in the chicken inner ear that may be involved in semicircular canal formation. *Dev Biol*. 1997;191(2): 215–29. <https://doi.org/10.1006/dbio.1997.8716>.
36. Wang W, Lufkin T. Hmx homeobox gene function in inner ear and nervous system cell-type specification and development. *Exp Cell Res*. 2005;306(2): 373–9. <https://doi.org/10.1016/j.yexcr.2005.03.016>.
37. Maroon H, Walshe J, Mahmood R, Kiefer P, Dickson C, Mason I. Fgf3 and Fgf8 are required together for formation of the otic placode and vesicle. *Development*. 2002;129(9):2099–108. <https://doi.org/10.1242/dev.129.9.2099>.
38. Feng Y, Xu Q. Pivotal role of hmx2 and hmx3 in zebrafish inner ear and lateral line development. *Dev Biol*. 2010;339(2):507–18. <https://doi.org/10.1016/j.ydbio.2009.12.028>.
39. Grimley E, Liao C, Ranghini EJ, Nikolovska-Coleska Z, Dressler GR. Inhibition of Pax2 transcription activation with a small molecule that targets the DNA binding domain. *ACS Chem Biol*. 2017;12(3):724–34. <https://doi.org/10.1021/acscchembio.6b00782>.
40. Doberstein K, Pfeilschifter J, Gutwein P. The transcription factor PAX2 regulates ADAM10 expression in renal cell carcinoma. *Carcinogenesis*. 2011; 32(11):1713–23. <https://doi.org/10.1093/carcin/bgr195>.
41. Fang LZ, Cai WT, Liu SL, Canela-Xandri O, Gao YH, Jiang JC, et al. Comprehensive analyses of 723 transcriptomes enhance genetic and biological interpretations for complex traits in cattle. *Genome Res*. 2020; 30(5):790–801. <https://doi.org/10.1101/gr.250704.119>.
42. Kelsey JA, Corl BA, Collier RJ, Bauman DE. The effect of breed, parity, and stage of lactation on conjugated linoleic acid (CLA) in milk fat from dairy cows. *J Dairy Sci*. 2003;86(8):2588–97. [https://doi.org/10.3168/jds.S0022-0302\(03\)73854-5](https://doi.org/10.3168/jds.S0022-0302(03)73854-5).
43. Falconer DS, Mackay TFC. Introduction to quantitative genetics, vol.1, Ed.4. Malaysia: Pearson Education Limited; 1996.

44. Kumar P, Henikoff S, Ng PC. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nat Protoc.* 2009;4(7):1073–81. <https://doi.org/10.1038/nprot.2009.86>.
45. Choi Y, Sims GE, Murphy S, Miller JR, Chan AP. Predicting the functional effect of amino acid substitutions and indels. *PLoS One.* 2012;7(10):e46688. <https://doi.org/10.1371/journal.pone.0046688>.

### **Publisher's Note**

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Ready to submit your research? Choose BMC and benefit from:**

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

**At BMC, research is always in progress.**

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

