Tohoku J. Exp. Med., 2015, **235**, 261-266

The CAA Repeat Polymorphism in the *ZFHX3* Gene Is Associated with Risk of Coronary Heart Disease in a Chinese Population

Shunchang Sun,[1] Wenwu Zhang,[2] Xi Chen[1] and Huiwen Song[1]

[1]Central Laboratory, Shenzhen Baoan Hospital, Southern Medical University, Shenzhen, Guangdong, P.R. China
[2]Department of Emergency Medicine, Shenzhen Baoan Hospital, Southern Medical University, Shenzhen, Guangdong, P.R. China

Coronary heart disease (CHD) is a disease resulting from the interaction between genetic variations and environmental factors. Zinc finger homeobox 3 (ZFHX3) is a transcription factor and contains a poly-glutamine tract in a compositionally biased region that is encoded by exon 9, containing a cluster of CAG and CAA triplets followed by the polymorphic CAA repeats: $(CAG)_2(CAA)_2(CAG)_3CAACAG(CAA)_nGCA$. Thus, nine successive glutamine residues precede the poly-glutamine tract, encoded by the polymorphic CAA repeats. The aim of this study was to investigate the association of the CAA repeat polymorphism in exon 9 of the *ZFHX3* gene with the risk of CHD in a Chinese population. The CAA repeat polymorphism was determined by polymerase chain reaction followed by DNA sequencing in 321 CHD patients. Genotype frequencies were compared using the non-parametric mood median test. Four alleles of $CAG(CAA)_{10}GCA$, $CAG(CAA)_8GCA$, $CAG(CAA)_9GCA$, and $CAG(CAA)_{11}GCA$ were found in Chinese CHD patients in exon 9 of the *ZFHX3* gene. The $CAG(CAA)_{10}GCA$ was a major allele (95.95%), and the $CAG(CAA)_8GCA$ was a minor allele (3.58%). The $CAG(CAA)_9GCA$ and $CAG(CAA)_{11}GCA$ were rare alleles (0.31% and 0.16%). The $CAG(CAA)_{10}GCA$ allele encodes a poly-glutamine tract of 19 residues. Importantly, the CHD patients homozygous for the $CAG(CAA)_{10}GCA$ allele had a higher risk of CHD, compared to the heterozygous patients carrying a $CAG(CAA)_8GCA$ allele. Moreover, the $CAG(CAA)_{10}GCA$ allele was significantly associated with hypertension, diabetes mellitus, or dyslipidemia ($P < 0.05$). Thus, the CAA repeat polymorphism in exon 9 of the *ZFHX3* gene contributes to the CHD susceptibility in the Chinese population.

Keywords: coronary heart disease; genotype; polymorphism; risk factor; zinc finger homeobox 3 gene
Tohoku J. Exp. Med., 2015 April, **235** (4), 261-266.

Introduction

Coronary heart disease (CHD) is a leading cause of morbidity and mortality in the world. It has been well established that genetic and environmental factors, such as CHD family history, hypertension, diabetes mellitus, dyslipidemia, poor diet, advanced age, and smoking habit, are associated with an increased risk of CHD (Lusk et al. 2014). Among these factors, hypertension, diabetes mellitus, and dyslipidemia are known to have a major influence. Although the exact genetic mechanism is unclear, genetic variations are estimated to account for about 30~60% of the CHD risk (Miller et al. 2014). Zinc finger homeobox 3 (ZFHX3) is a transcriptional repressor of 3,703 amino acids containing one ATPase A-motif, two DEAH box-like sequences, four homeodomains, and 23 zinc finger motifs involved in transcriptional regulation (Fig. 1) (Dong et al. 2010). ZFHX3 inhibits the enhancer of the alpha-fetoprotein (AFP) gene by binding to its AT-rich core sequence, and it also regulates myoblasts differentiation through the binding to the AT-rich sequence of *MYF6* promoter (Li et al. 2013). The ZFHX3 contains a polymorphic poly-glutamine tract that is encoded by a trinucleotide CAA repeat in a compositionally biased region (Benjamin et al. 2009). The compositionally biased region is a stretch in protein sequences made from mainly a distinct subset of amino acid residues. Such a region is frequently linked to a structural role. The ZFHX3 contains 15 compositionally biased regions. The *ZFHX3* gene is located on chromosome 16q22.3, spanning at least 265 kb of genomic DNA with 10 exons (Fig. 2), and the size of a predicted cDNA is around 16 kb (Jiang et al. 2014). Alternative splicing is involved in the generation of the two ZFHX3 isoforms, and the splicing variant is associated with neuronal differentiation (Jung et al. 2005). Genetic variants in the *ZFHX3* gene are associated with atrial fibrillation in individuals of European

Received January 9, 2015; revised and accepted March 2, 2015. Published online March 21, 2015; doi: 10.1620/tjem.235.261.
Correspondence: Shunchang Sun, Central Laboratory, Shenzhen Baoan Hospital, Southern Medical University, 118 Longjing Er Road, Baoan, Shenzhen, Guangdong 518101, P.R. China.
e-mail: shunchangsun@aliyun.com

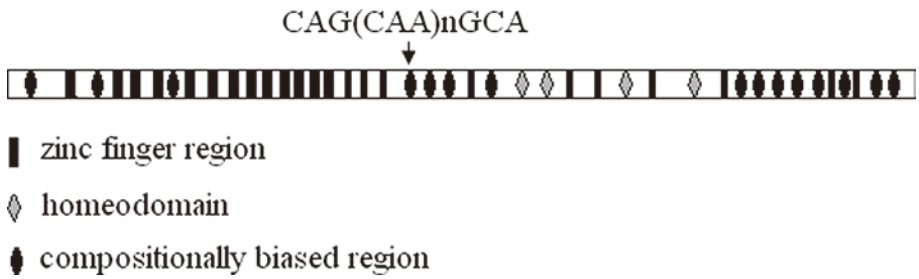

Fig. 1. Schematic diagram of the ZFHX3 protein.
The ZFHX3 protein has three types of domain: zinc finger region, homeodomain, and compositionally biased region.

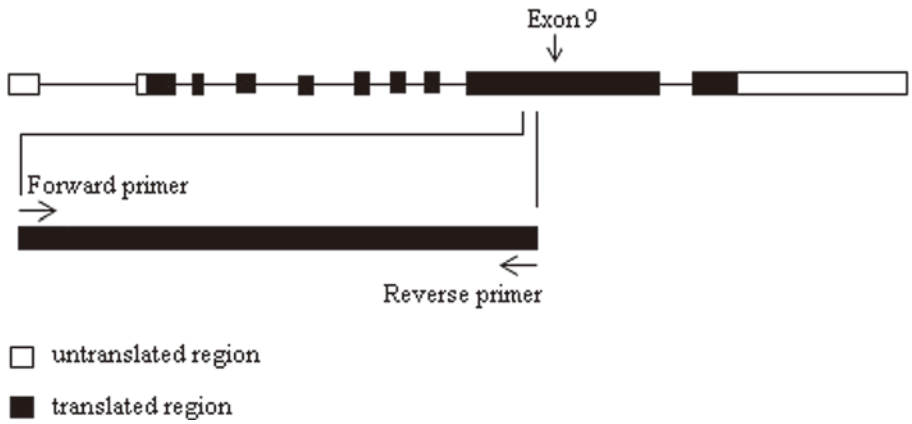

Fig. 2. Diagrammatic representation of the *ZFHX3* gene.

ancestry and in a Chinese Han population (Li et al. 2011). There was still no research data on the relation between the CAA repeat polymorphism encoding a polyglutamine tract in the fourth compositionally biased region of the ZFHX3 and CHD.

Here, we aimed to reveal the relationship between the CAA repeat polymorphism and CHD in a Chinese population. We subsequently found that the CAG(CAA)$_{10}$GCA allele was a risk factor for CHD in the Chinese population.

Methods

Subjects

A total of 321 hospitalized CHD patients were recruited during December 2011 to August 2013 from Department of Cardiology, Shenzhen Baoan Hospital, Southern Medical University. All CHD patients were diagnosed according to ischemic heart disease diagnostic criteria issued by WHO in 1979 or the presence of stenosis of more than 50% luminal diameter in at least one significant coronary artery on coronary angiography (Patel et al. 2014). Patients with coronary artery bypass surgery were also considered as CHD cases. All these CHD patients were unrelated Chinese Han people. Baseline clinical characteristics of CHD patients enrolled in this study are summarized in Table 1. This study was approved by the ethics committee of Shenzhen Baoan Hospital, Southern Medical University, and was kept in accordance with the Helsinki Declaration of 1975, as revised in 1983. Informed written consent was obtained from all CHD patients.

Table 1. Baseline clinical characteristics of CHD patients (n = 321).

Characteristics	
Gender (male/female)	257/64
Age	58.8 ± 13.9
Systolic blood pressure (mmHg)	130.3 ± 24.7
Diastolic blood pressure (mmHg)	75.6 ± 16.6
Body mass index	25.3 ± 3.1
Fasting plasma glucose (mmol/L)	7.7 ± 2.7
Total cholesterol (mmol/L)	4.8 ± 1.3
Triglycerides (mmol/L)	1.6 ± 1.2
Smoking (past or current)	42.4%
Drinking	8.4%
Hypertension (past history)	48.3%
Diabetes mellitus (past history)	26.2%
Dyslipidemia (past history)	63.9%

Polymorphism analysis

Genomic DNA was extracted from peripheral blood lymphocytes using a phenol-chloroform extraction method after proteinase K digestion for all CHD patients (Miller et al. 1988), and it was quantified using a spectrophotometer. An absorbance ratio of 1.8 : 2.0 or greater was considered acceptable and the final DNA solution was stored at −70°C. The sequences encoding the fourth compositionally biased region of the ZFHX3 was amplified by polymerase chain reaction (PCR) using the following primers: 5′-ctctttggcgtttcttgctg-3′ (for-

ward) and 5′-cctgaagagccagaagcaga-3′ (reverse) (Fig. 2). PCR was performed using 100 ng of genomic DNA in 50 μl reactions containing 10 mM Tris-HCl, pH 8.3, 50 mM KCl, 2.5 mM $MgCl_2$, 200 μM of each dNTP, 2 U Taq polymerase, and 0.5 μM each primer. Following denaturation at 94°C for 5 minutes, amplification was performed using 30 cycles at 94°C for 50 seconds, annealing for 45 seconds at 61°C, and 72°C for 50 seconds. Amplification products were visualized on 2% agarose gel. Amplified PCR fragments were subjected to bidirectional direct and cycle sequencing after purification. Direct sequencing was performed using the BigDye™ Terminator on ABI3730 Genetic Analyzer. Cycle sequencing was performed using Stratagene's RoboCycler Gradient 96 temperature cycler with Hot Top Assembly.

Statistical analysis

Statistical analysis was performed using SPSS Statistics 13.0 for Windows. Allele frequencies were compared using the non-parametric mood median test. Differences among categorical variables were analyzed using Pearson's chi-square test. A P value < 0.05 was considered statistically significant.

Results

Four alleles of $CAG(CAA)_{10}GCA$, $CAG(CAA)_8GCA$, $CAG(CAA)_9GCA$, and $CAG(CAA)_{11}GCA$ were found in exon 9 of the *ZFHX3* gene encoding the fourth compositionally biased region in CHD patients (Fig. 3). The $CAG(CAA)_{10}GCA$ was a major allele, and most patients were homozygous for this allele (Table 2). The $CAG(CAA)_8GCA$ was a minor allele, while the $CAG(CAA)_9GCA$ and $CAG(CAA)_{11}GCA$ were rare alleles. Thus, minor and rare alleles were identified in a heterozygous state in 26 CHD patients (Table 2). The frequency of each allele is shown in Table 3 for CHD patients.

We also showed that CHD patients who are homozygous for the $CAG(CAA)_{10}GCA$ allele were more likely to develop hypertension, diabetes mellitus, or dyslipidemia compared to the heterozygous patients carrying the $CAG(CAA)_{10}GCA$ and $CAG(CAA)_8GCA$ alleles (Table 4). Thus, the $CAG(CAA)_{10}GCA$ allele was significantly associated with hypertension, diabetes mellitus, or dyslipidemia in the Chinese CHD patients ($P < 0.05$). In this study, however, we were unable to analyze whether either $CAG(CAA)_{11}GCA$ or $CAG(CAA)_9GCA$ allele is associated with hypertension, diabetes mellitus, or dyslipidemia because of low allele frequencies and inadequate subject numbers.

Discussion

Trinucleotide repeats belong to a family of microsatel-

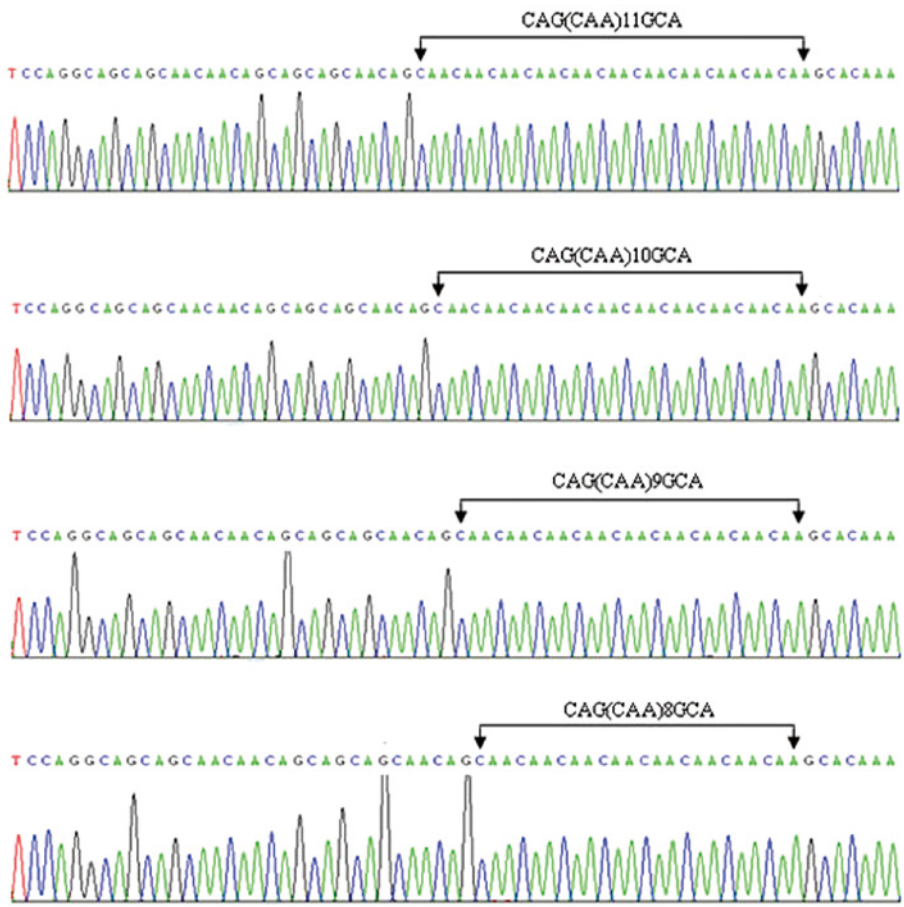

Fig. 3. The $CAG(CAA)_nGCA$ polymorphism encoding the fourth compositionally biased region in the *ZFHX3* gene in CHD patients.

Table 2. The $CAG(CAA)_nGCA$ genotype in CHD patients.

$CAG(CAA)_nGCA$ genotype	CHD patients (*n*)
$CAG\ (CAA)_{10}GCA/CAG\ (CAA)_{10}GCA$	295
$CAG\ (CAA)_{10}GCA/CAG\ (CAA)_{8}GCA$	23
$CAG\ (CAA)_{10}GCA/CAG\ (CAA)_{9}GCA$	2
$CAG\ (CAA)_{10}GCA/CAG\ (CAA)_{11}GCA$	1

Table 3. Frequency of the $CAG(CAA)_nGCA$ allele of the *ZFHX3* gene in CHD patients.

$CAG(CAA)_nGCA$ allele	Poly-Gln	Frequency
$(CAG)_2(CAA)_2(CAG)_3CAACAG(CAA)_{10}GCA$	$(Gln)_{19}$	95.95%
$(CAG)_2(CAA)_2(CAG)_3CAACAG(CAA)_{8}GCA$	$(Gln)_{17}$	3.58%
$(CAG)_2(CAA)_2(CAG)_3CAACAG(CAA)_{9}GCA$	$(Gln)_{18}$	0.31%
$(CAG)_2(CAA)_2(CAG)_3CAACAG(CAA)_{11}GCA$	$(Gln)_{20}$	0.16%

Table 4. Frequency of the $CAG(CAA)_nGCA$ genotype in CHD patients with hypertension, diabetes mellitus, or dyslipidemia.

Genotype	CHD with hypertension	CHD with diabetes mellitus	CHD with dyslipidemia
$CAG(CAA)_{10}GCA/\ CAG(CAA)_{10}GCA$	50.51% (149/295)	27.46% (81/295)	65.42% (193/295)
$CAG(CAA)_{8}GCA/\ CAG(CAA)_{10}GCA$	21.74% (5/23)	8.70% (2/23)	43.48% (10/23)
χ^2	7.071	3.894	4.451
P	< 0.01	< 0.05	< 0.05

lite sequences, also known as short tandem repeats. Tandem repeats of different trinucleotide motifs are present in the human transcriptome, and their functions may depend on the structures they form (Malgowska et al. 2014). Some AT-rich trinucleotide repeat types, such as CAA and CTT, are particularly underrepresented in exons, whereas GC-rich repeats (CGG and CCG) are highly overrepresented, implying that trinucleotide repeat sequences have a functional significance (Naumann et al. 2014). A survey of the human genome reference sequence revealed that it harbors more than 32,000 tracts of trinucleotide repeat sequences composed of six or more repeated units (Axford et al. 2013). In human exons, which account for less than 3% of the genomic sequence, there are as many as 1,030 trinucleotide-repeat tracts (Slean et al. 2013). Polymorphic trinucleotide repeats are better tolerated than dinucleotide and tetranucleotide repeats in translated sequences because their length variation does not change the open reading frame (Panigrahi et al. 2012). About 60% of exonic trinucleotide repeats are primarily translated to a poly-Gln, poly-Ala, poly-Glu, or poly-Leu tract (Kozlowski et al. 2010). A study showed that the amino acid-coding property of trinucleotide repeats is not the only feature for which these sequences are selected in exons (Sobczak et al. 2010). The other properties of trinucleotide repeat sequences that manifest themselves on the levels of DNA, RNA, or both may also contribute to the functional importance of these sequences and their prevalence in exons (Sobczak et al. 2010). These properties of trinucleotide repeats may include their ability to form higher order structures in single-stranded DNA and transcripts (Liu et al. 2012). Higher order structures may play an important regulatory role in numerous cellular processes, such as DNA replication repair and at various steps of gene expression ranging from transcription to mRNA decay (Gannon et al. 2012).

The *ZFHX3* gene, also called AT motif-binding factor 1 (ATBF1), was first described as a transcription factor that inhibits the human *AFP* gene expression in the liver (Sun et al. 2012). It has been reported to be a tumor suppressor gene in multiple cancers (Sun et al. 2014). The ZFHX3 is a DNA-binding protein that contains multiple homeodomains and zinc finger motifs (Dong et al. 2012). The *ZFHX3* gene has been associated with growth and differentiation regulation of several tissues, such as neuron and skeletal muscle (Jung et al. 2005). Although the function of the *ZFHX3* gene in cardiac tissue is unknown, it was expressed in mouse hearts (Dong et al. 2011). The *ZFHX3* gene variants are associated with atrial fibrillation in a Chinese Han population (Sun et al. 2014).

CHD is the most common form of cardiovascular disease with high morbidity and mortality. In this study, we identified the polymorphic CAA repeat, $CAG(CAA)_nGCA$, in exon 9 of the *ZFHX3* gene that encodes the fourth compositionally biased region in CHD patients and healthy subjects in a Chinese population. The $CAG(CAA)_nGCA$ polymorphism in the *ZFHX3* gene here described has not been reported previously. The CAA repeat polymorphism encodes a poly-Gln tract (17-20 Gln residues) in the

ZFHX3 protein. Four alleles of $CAG(CAA)_{10}GCA$, $CAG(CAA)_8GCA$, $CAG(CAA)_9GCA$, and $CAG(CAA)_{11}GCA$ were found in exon 9 of the *ZFHX3* gene in CHD patients. Several lines of evidence point to the critical involvement of poly-Gln aggregation in the disease process (Stork et al. 2005). Aggregation kinetics of proteins containing a poly-Gln tract exhibits dependency on repeat length that qualitatively mirrors the repeat length dependency of disease risk (Sobczak et al. 2010). The frequency of the $CAG(CAA)_{10}GCA$ allele that encodes the ZFHX3 protein containing $(Gln)_{19}$ was 95.95% (Table 3). In a healthy Chinese population, the $CAG(CAA)_{10}GCA$ and $CAG(CAA)_8GCA$ were wild-type alleles, and their frequencies were 96.17% and 3.83% (unpublished data). There was no significant difference in two allele frequencies between CHD patients and healthy controls. Among CHD patients, the $CAG(CAA)_8GCA$ was a minor allele. The $CAG(CAA)_9GCA$ and $CAG(CAA)_{11}GCA$ were rare alleles, but statistical analysis was not performed for frequency difference between them. Little is known about the physiological functions of the poly-Gln tract in the ZFHX3 protein.

In this study, we showed that the $CAG(CAA)_{10}GCA$ allele was significantly associated with hypertension, diabetes mellitus, or dyslipidemia compared to the $CAG(CAA)_8GCA$ allele in CHD patients. Hypertension, diabetes mellitus, and dyslipidemia are major risk factors for CHD (Chen et al. 2014); therefore, individuals carrying the $CAG(CAA)_{10}GCA$ allele are likely to develop CHD. Expansion of CAA repeats encoding the poly-Gln tract in the fourth compositionally biased region of ZFHX3 may be a risk factor for CHD in the Chinese population. More studies should be focused on the functions of the compositionally biased region of the ZFHX3.

In summary, four alleles of $CAG(CAA)_{10}GCA$, $CAG(CAA)_8GCA$, $CAG(CAA)_9GCA$, and $CAG(CAA)_{11}GCA$ were found in exon 9 of the *ZFHX3* gene encoding the fourth compositionally biased region in Chinese CHD patients. This study also showed that the $CAG(CAA)_{10}GCA$ allele was significantly associated with hypertension, diabetes mellitus, or dyslipidemia in CHD patients, compared to the $CAG(CAA)_8GCA$ allele. Thus, expansion of the CAA repeats, encoding the poly-Gln tract in the fourth compositionally biased region, in the *ZFHX3* gene may be a risk factor for CHD in the Chinese population. Further investigation is needed to clarify the functions of the compositionally biased region of ZFHX3.

Authors' Contribution

Sun SC conceived experiment design, and drafted and revised manuscript. Zhang WW participated in sample collection and analysis of data. Chen X carried out PCR and data analysis. Song HW performed DNA sequencing. All authors read and approved the final manuscript.

Acknowledgments

This study was supported by research grants from Science and Technology Innovation Commission of Shenzhen Municipality (No. JCYJ20140414103602832).

Conflict of Interest

The authors declare no conflict of interest.

References

Axford, M.M., Wang, Y.H., Nakamori, M., Zannis-Hadjopoulos, M., Thornton, C.A. & Pearson, C.E. (2013) Detection of slipped-DNAs at the trinucleotide repeats of the myotonic dystrophy type I disease locus in patient tissues. *PLoS Genet*., **9**, e1003866.

Benjamin, E.J., Rice, K.M., Arking, D.E., Pfeufer, A., van Noord, C., Smith, A.V., Schnabel, R.B., Bis, J.C., Boerwinkle, E., Sinner, M.F., Dehghan, A., Lubitz, S.A., D'Agostino, R.B. Sr., Lumley, T., Ehret, G.B., et al. (2009) Variants in ZFHX3 are associated with atrial fibrillation in individuals of European ancestry. *Nat. Genet*., **41**, 879-881.

Chen, S., Xie, X., Wang, Y., Gao, Y., Xie, X., Yang, J. & Ye, J. (2014) Association between leukocyte mitochondrial DNA content and risk of coronary heart disease: a case-control study. *Atherosclerosis*, **237**, 220-226.

Dong, X.Y., Fu, X., Fan, S., Guo, P., Su, D. & Dong, J.T. (2012) Oestrogen causes ATBF1 protein degradation through the oestrogen-responsive E3 ubiquitin ligase EFP. *Biochem. J*., **444**, 581-590.

Dong, X.Y., Guo, P., Sun, X., Li, Q. & Dong, J.T. (2011) Estrogen up-regulates ATBF1 transcription but causes its protein degradation in estrogen receptor-alpha-positive breast cancer cells. *J. Biol. Chem*., **286**, 13879-13890.

Dong, X.Y., Sun, X., Guo, P., Li, Q., Sasahara, M., Ishii, Y. & Dong, J.T. (2010) ATBF1 inhibits estrogen receptor (ER) function by selectively competing with AIB1 for binding to the ER in ER-positive breast cancer cells. *J. Biol. Chem*., **285**, 32801-32809.

Gannon, A.M., Frizzell, A., Healy, E. & Lahue, R.S. (2012) MutSβ and histone deacetylase complexes promote expansions of trinucleotide repeats in human cells. *Nucleic Acids Res*., **40**, 10324-10333.

Jiang, Q., Ni, B., Shi, J., Han, Z., Qi, R., Xu, W., Wang, D., Wang, D.W. & Chen, M. (2014) Down-regulation of ATBF1 activates STAT3 signaling via PIAS3 in pacing-induced HL-1 atrial myocytes. *Biochem. Biophys. Res. Commun*., **449**, 278-283.

Jung, C.G., Kim, H.J., Kawaguchi, M., Khanna, K.K., Hida, H., Asai, K., Nishino, H. & Miura, Y. (2005) Homeotic factor ATBF1 induces the cell cycle arrest associated with neuronal differentiation. *Development*, **132**, 5137-5145.

Kozlowski, P., Sobczak, K. & Krzyzosiak, W.J. (2010) Trinucleotide repeats: triggers for genomic disorders? *Genome Med*., **2**, 29.

Li, C., Wang, F., Yang, Y., Fu, F., Xu, C., Shi, L., Li, S., Xia, Y., Wu, G., Cheng, X., Liu, H., Wang, C., Wang, P., Hao, J., Ke, Y., et al. (2011) Significant association of SNP rs2106261 in the ZFHX3 gene with atrial fibrillation in a Chinese Han GeneID population. *Hum. Genet*., **129**, 239-246.

Li, M., Zhao, D., Ma, G., Zhang, B., Fu, X., Zhu, Z., Fu, L., Sun, X. & Dong, J.T. (2013) Upregulation of ATBF1 by progesterone-PR signaling and its functional implication in mammary epithelial cells. *Biochem. Biophys. Res. Commun*., **430**, 358-363.

Liu, C.R., Chang, C.R., Chern, Y., Wang, T.H., Hsieh, W.C., Shen, W.C., Chang, C.Y., Chu, I.C., Deng, N., Cohen, S.N. & Cheng, T.H. (2012) Spt4 is selectively required for transcription of extended trinucleotide repeats. *Cell*, **148**, 690-701.

Lusk, C.M., Dyson, G., Clark, A.G., Ballantyne, C.M., Frikke-Schmidt, R., Tybjærg-Hansen, A., Boerwinkle, E. & Sing, C.F.

(2014) Validated context-dependent associations of coronary heart disease risk with genotype variation in the chromosome 9p21 region: the Atherosclerosis Risk in Communities study. *Hum. Genet*., **133**, 1105-1116.

Malgowska, M., Gudanis, D., Kierzek, R., Wyszko, E., Gabelica, V. & Gdaniec, Z. (2014) Distinctive structural motifs of RNA G-quadruplexes composed of AGG, CGG and UGG trinucleotide repeats. *Nucleic Acids Res*., **42**, 10196-10207.

Miller, C.L., Haas, U., Diaz, R., Leeper, N.J., Kundu, R.K., Patlolla, B., Assimes, T.L., Kaiser, F.J., Perisic, L., Hedin, U., Maegdefessel, L., Schunkert, H., Erdmann, J., Quertermous, T. & Sczakiel, G. (2014) Coronary heart disease-associated variation in TCF21 disrupts a miR-224 binding site and miRNA-mediated regulation. *PLoS Genet*., **10**, e1004263.

Miller, S.A., Dykes, D.D. & Polesky, H.F. (1988) A simple salting out procedure for extracting DNA from human nucleated cells. *Necleic Acids Res*., **16**, 1215.

Naumann, A., Kraus, C., Hoogeveen, A., Ramirez, C.M. & Doerfler, W. (2014) Stable DNA methylation boundaries and expanded trinucleotide repeats: role of DNA insertions. *J. Mol. Biol*., **426**, 2554-2566.

Panigrahi, G.B., Slean, M.M., Simard, J.P. & Pearson, C.E. (2012) Human mismatch repair protein hMutLα is required to repair short slipped-DNAs of trinucleotide repeats. *J. Biol. Chem*., **287**, 41844-41850.

Patel, R.S., Asselbergs, F.W., Quyyumi, A.A., Palmer, T.M., Finan, C.I., Tragante, V., Deanfield, J., Hemingway, H., Hingorani, A.D. & Holmes, M.V. (2014) Genetic variants at chromosome 9p21 and risk of first versus subsequent coronary heart disease events: a systematic review and meta-analysis. *J. Am. Coll. Cardiol*., **63**, 2234-2245.

Slean, M.M., Reddy, K., Wu, B., Nichol Edamura, K., Kekis, M., Nelissen, F.H., Aspers, R.L., Tessari, M., Schärer, O.D., Wijmenga, S.S. & Pearson, C.E. (2013) Interconverting conformations of slipped-DNA junctions formed by trinucleotide repeats affect repair outcome. *Biochemistry*, **52**, 773-785.

Sobczak, K., Michlewski, G., de Mezer, M., Kierzek, E., Krol, J., Olejniczak, M., Kierzek, R. & Krzyzosiak, W.J. (2010) Structural diversity of triplet repeat RNAs. *J. Biol. Chem*., **285**, 12755-12764.

Stork, M., Giese, A., Kretzschmar, H.A. & Tavan, P. (2005) Molecular dynamics simulations indicate a possible role of parallel beta-helices in seeded aggregation of poly-Gln. *Biophys. J*., **88**, 2442-2451.

Sun, X., Fu, X., Li, J., Xing, C., Frierson, H.F., Wu, H., Ding, X., Ju, T., Cummings, R.D. & Dong, J.T. (2014) Deletion of atbf1/zfhx3 in mouse prostate causes neoplastic lesions, likely by attenuation of membrane and secretory proteins and multiple signaling pathways. *Neoplasia*, **16**, 377-389.

Sun, X., Fu, X., Li, J., Xing, C., Martin, D.W., Zhang, H.H., Chen, Z. & Dong, J.T. (2012) Heterozygous deletion of Atbf1 by the Cre-loxP system in mice causes preweaning mortality. *Genesis*, **50**, 819-827.