

The RAB38-Hosted Chicken *Mir-1657* Gene: Sequence Conservation, Seed Region Polymorphisms and Putative Targets

¹Chuan-Sheng Zhang, ²Li-Ying Geng, ²Yang-Qing, ¹Wen-Jin Zhu, ¹Qiu-Yue Wang, ¹Xiang-Zhai Zhang, ¹Hong-Nuan Sun, ¹Xiao-Song Li, ¹Rui-Zao Liu and ¹Xiao-Hui Liu
¹Department of Animal Science, ²Department of Life Science,
Hebei Normal University of Science and Technology, Qinhuangdao, 066600 Changli, P.R. China

Abstract: The aim of this study was first to investigate the evolutionary conservatism of *mir-1657* among in different species and then to investigated a G-A polymorphism in the seed region of mature *gga-mir-1657* sequence with forced PCR-RFLP using *Ava* III nuclease in six chicken populations; 178 individuals from Beijing Fatty chicken (BF), Wenshangluhua chicken (WL), Jiningbairi chicken (JB), Siyuwugu chicken (WG) and Langya chicken (LY) breeds. Bioinformatics analyses indicated that *gga-mir-1657* gene G>A polymorphism may alter target selection and secondary structure. The findings indicate that the rs14934924 SNP may exert profound biological effects in the formation of some special phenotype of chicken and enables functional annotation of *gga-mir-1657*.

Key words: *Gga-mir-1657*, chicken, SNP, functional annotation, biological, China

INTRODUCTION

Mir-1657 gene was 1st reported in chicken which is located in the intron 2 of the *RAB38* host gene on 1 chromosome (Glazov *et al.*, 2008). *RAB38* is a member of RAS oncogene family which encodes a small G-protein involved in endoplasmic reticulum-related vesicle transport, generates specific antibody and T cell responses in melanoma patients and is highly expressed in melanoma tissue and absent in normal tissue. Recently, the intronic microRNAs, transcribed together with the host gene *mRNA* were found may serve the interest of its host gene by silencing a cohort of genes that are functionally antagonistic to the host gene itself (Lutter *et al.*, 2010). However, little is known about the relationship between the target genes of *gga-mir-1657* and *RAB38* in chickens.

Many evidences indicated that SNPs located the seed regions (positions 2-7 from the 5'end of the mature microRNA) may modify various biological processes by influencing the processing and/or target selection of microRNAs having long ranging phenotypic effects (Mishra *et al.*, 2008). There is a rapidly growing interest for SNPs in seed region of microRNA in genetic analyses as several studies have suggested association between microRNA-seed-SNPs and human cancer risk (Ryan *et al.*, 2010). The previous studies demonstrated that a G>A polymorphisms (rs14934924 SNP) occurs in

seed regions of *gga-mir-1657* gene (Liyang *et al.*, 2009). This SNP has not been studied in details so far. Therefore, the objectives of the present study were to reveal the sequence conservation of *gga-mir-1657* among different species; elucidate the effect of the rs14934924 SNP on pre-investigate's secondary structure; investigate the distribution of the G>A polymorphisms of *gga-mir-1657* among six chicken populations and identify putative target genes of *gga-mir-1657*.

MATERIALS AND METHODS

Samples and DNA extraction: Blood samples were taken from 178 individuals of 5 indigenous Chinese chicken breeds and 1 imported chicken breed, Leghorn chicken (Table 1). DNA extraction was conducted by the phenol extraction method.

Primer design and PCR amplification: The fragment (210 bp) of *gga-mir-1657* (MI0007391) was amplified using primers designed by primer 3 (http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi) based on the chicken sequence from mirbase. The primers were forward 5'-TTCTGAAGGTGGACTTCATGG-3' and reverse 5'-CGTATCACACACCAACAATGC-3'. The PCR reaction volume of 25 μ L contained approximately 50 ng of genomic DNA, 1.25 mM Taq DNA polymerase, 2.5 μ L of 1 \times PCR buffer, 1.5 mM MgCl₂, 0.2 mM dNTP and 10 pM

Table 1: Chicken breeds sampled for this study

Breed (code)	Sample size	Location
Beijing Fatty chicken (BF)	29	Beijing
Wenshangluhua chicken (WL)	30	Shandong Wenshang
Jiningbairi chicken (JB)	29	Shandong Jining
Siyuwugu chicken (WG)	30	Beijing
Leghorn chicken (LH)	30	Beijing
Langya chicken (LY)	30	Shandong Yantai

of each primer. Amplification conditions included an initial denaturation at 94°C for 4 min followed by 35 cycles at 94°C for 30 sec, 60°C for 30 sec and 72°C for 30 sec followed by a final extension at 72°C for 10 min. PCR products were detected on 3.5% agarose gel.

PCR-RFLP and genotype determination: The gga-mir-1657 PCR product was digested with 10 units of Ava III restriction enzyme and 10 µL of PCR product at 37°C overnight in a water bath. The digested products were detected by electrophoresis in 3.5% agarose gel stained with Ethidium Bromide (EB). The homozygote AA was defined when base A exists at position 5 of seed region forming ATGCAT which could be recognized by Ava III, producing fragments of 190 and 20 bp (the 20 bp could not be seen in the gel). GG was defined when A was replaced by G which could not be recognized by Ava III; only a 190 bp fragment appeared on the gel. The heterozygote AG was defined when G and A existed simultaneously at the homologous chromosome, forming fragments of 210, 190 and 20 bp in the gel.

Secondary structure alterations of variant gga-mir-1657 precursors: The most stable secondary RNA structure with the lowest free energy for pre-gga-mir-1657 with A>G alleles were calculated using M-fold (Zuker, 2003). The absolute difference of free energy for pre-gga-mir-1657 with different alleles were used as the parameter for the assessment of the impact on secondary structure of pre-gga-mir-1657.

Impact of SNP on gga-mir-1657 target genes: The chicken Unigene (NCBI) was scanned for potential gga-mir-1657 targets using the miRanda algorithm (Version 3.1) (Enright *et al.*, 2003) with the default parameters for score threshold (>130) and free energy threshold (<-16). The predicted targets were further filtered using more stringent criteria in which they must contain either a match between nucleotides 2-8 of the microRNA with the target sequence or a match between nucleotides 2-8 of the microRNA with the target sequence (G:U base-pairing was not tolerated).

Statistical methods: The genetic diversity of A>G located in the seed region of gga-mir-1657 gene for each

population was estimated using *Pop gene 32* software (Yeh *et al.*, 1999) including observed heterozygosity, expected heterozygosity, effective number of alleles and Shannon's information index.

RESULTS AND DISCUSSION

Gene organization, comparative genomics and secondary structure alterations of variant gga-mir-1657 precursors. According to the miRBase 13.0 (<http://microna.sanger.ac.uk/>) and Ensembl (<http://www.ensembl.org/index.html>) databases, the gga-mir-1657 resides within the intron 2 of the *RAB38* gene (Fig. 1). The A>G located in the seed region of gga-mir-1657 could alter free energy values and alter the predicted RNA secondary structure with Mfold program (Fig. 1). It has previously been shown that recognition of the ssRNA-dsRNA junction and adjacent ~11 bp stem by DGCR8 is critical for the processing of pri-microRNA. Therefore, the SNPs occurring in mir-1657 may have a larger impact on the maturation of microRNA and have important phenotypic consequences. Recently, mir-146a-SNP (rs2910164) within the pre-miR-146a sequence reduced both the amount of pre-miR-146a and mature miR-146a and apparently affected the Drosha/DGCR8 processing step.

Orthologs of gga-mir-1657 were retrieved from the Ensembl genome browser using comparative genomics/alignments/19 amniota vertebrates pecan option. Multiple species sequence alignment was performed using the MultAlin program (<http://bioinfo.genotoul.fr/multalin>). Multiple alignment of known (chicken) and predicted gga-mir-1657 ortholog sequences showed a high level of conservation with *Meleagris gallopavo* and a low level of conservation with *Taeniopygia guttata* among 3 neognath birds EPO but absent in other species (Fig. 2).

The mir-1657 ortholog in *Meleagris gallopavo* showed the typical stem-loop secondary structure (Fig. 3a) but absent in *Taeniopygia guttata* (Fig. 3b). So, researchers speculated that the mir-1657 may be a galliformes-specific microRNA. These primate-specific microRNA families have been considered that may contribute to developmental novelties during evolution. As miRNAs are involved in gene regulations, the galliformes-specific mir-1657 may contribute to various biological processes and play critical roles in disease or population disease susceptibility in chicken.

Impact of SNP on gga-mir-1657 target genes: The SNP (rs14934924) is located in the crucial seed sequence of gga-mir-1657. So, it determines its complementarity to

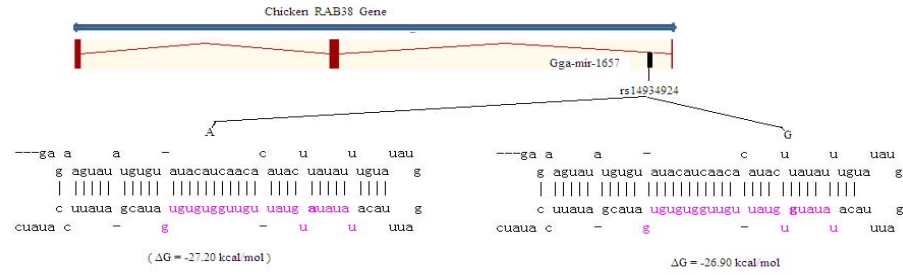


Fig. 1: Gene organization the *gga-mir-1657* gene and *RAB38* gene

Callus gallus	GAGAAGTATATGTTATACATCAACACATCTATAT--TTGTATGGATTTACATA--
Meleagris gallopavo	GAGAACTACGTATATACATCAGCACACATATAT--TTGTAGATGGATTTACATA--
Taeniopygia guttata	AAGTAGTGTGTCATATACACAAATGTATGCATATAT--ATTGGTATGCATTTGCATA--
Anolis carolinensis
mus_musculus
Rattus norvegicus
Oryzotolagus cuniculus
Pan troglodytes
Homo sapiens
Gorilla gorilla
Pongo pygmaeus
Macaca mulatta
Callithrix jacchus
Equus caballus
Canis lupus familiaris
Sus scrofa
Bos taurus
Monodelphis domestica
Ornithorhynchus anatinus
Callus gallus
Meleagris gallopavo
Taeniopygia guttata
Anolis carolinensis
mus_musculus
Rattus norvegicus
Oryzotolagus cuniculus
Pan troglodytes
Homo sapiens
Gorilla gorilla
Pongo pygmaeus
Macaca mulatta
Callithrix jacchus
Equus caballus
Canis lupus familiaris
Sus scrofa
Bos taurus
Monodelphis domestica
Ornithorhynchus anatinus

Fig. 2: Comparative genomic of the *gga-mir-1657* gene

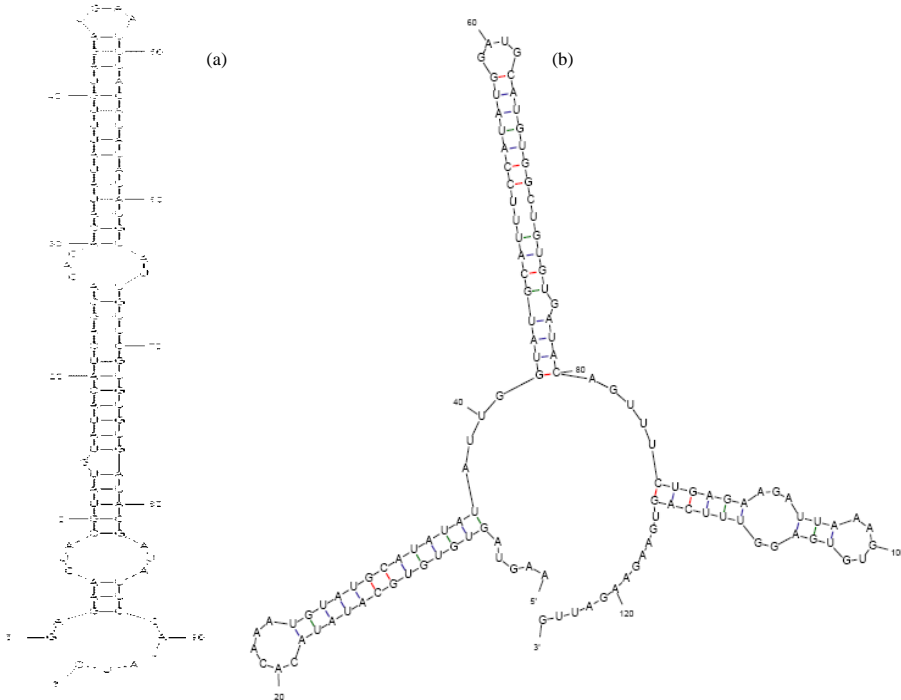


Fig. 3: a) Secondary structure of *gga-mir-1657* ortholog in *Meleagris gallopavo* and b) *Taeniopygia guttata*

Table 2: Continued

Genes	gga-mir-1657-A	kcal/mol
IRF4	3' GTGTGTGG-TTGTATGTGTATAT 5' ::: :: 5' TGTATATTGAATTTTGCACATATT 3'	-16.44
TMEM9B	3' GTGTGTGG-TT-GTT-AT-GTGTATAT 5' ::: : 5' TGTACATCAAATTCATTACCACATATG 3'	-17.67
Chkmm2II	3' GTGTGTGGTTGT-TATGTGTATAT 5' ::: : 5' TGTA-GCTGATACATACACATATA 3'	-17.62
EMB	3' GT-GTGTGGTTTGT-TATGTGTATAT 5' : :: 5' CATTTTCATCTATAATGCCACATATG 3'	-17.79
CD31	3' GTG-T-G-TGG-T-TGT-TATGTGTATAT 5' : :: 5' CACTGATACCTAAACA-ACACATATT 3'	-14.17
2i2i	3' G-TGT-GTGGT-TGT-TATGTGTATAT 5' : : :: 5' CTGCATCA-GGTGCAGGTCCACATATT 3'	-19.02
33β	3' GTG-TGTGGTTG-TTATGTGTATAT 5' : : 5' TATGAAATC-ACTGATACACATATA 3'	-16.32
MFAP5	3' GTG-TG-TGGT-TGT-TATGTGTATAT 5' : :: 5' CACTACTATTGTAC-AT-CACATATA 3'	-14.23
CD164	3' G-TGTGTG-GT-TGT-T-ATGTGTATAT 5' : : :: 5' CGGTACGTGTGTGTACTGCACATATA 3'	-16.8
SSEBP2	3' GT-GTGTG-G-T-TG-TT-AT-GTGTATAT 5' : :: 5' CATTGCTCTTTCACAAAGTACCACATATT 3'	-19.83
GLO1	3' GTGTGT-GG-TTGTATGTGTATAT 5' : : 5' TACATATCCTTGTG-TACACATATT 3'	-16.78
21h11	3' G-TG-TG-TGGTGT-TATGTGTATAT 5' : : : 5' CTGCTACTGCT-GCCATCCACATATA 3'	-19.72
1.9e21	3' GTG-TG-T-GGTT-G-T-T-ATGTGTATAT 5' : : 5' C-CTGTGATCTGGTTTCCACTACACATATA 3'	-14.1
EMP1	3' G-TGTGT-GGTT-GTTATGTGTATAT 5' : : : 5' CTACTTAGCCAACCAAT-CACATATG 3'	-21.26
PL15	3' GTG-T-GTGGTTGT-TATGTATATAT 5' : : : 5' CATGAGTAGC-ATGTATGCATATATA 3'	-15.15
FGF19	3' GTGTGTGGT-TGT-TATGTATATAT 5' : : 5' GAC-TG-CATATATGTACATATATA 3'	-18
SPRY1	3' GTGTGTG-GTTGTT-ATGTATATAT 5' : : 5' AGCATATGGAA-AATTACATATATT 3'	-17.09
ESR1	3' GTGTG-T-GG-TTGTATGTATATAT 5' : 5' CATACAAAGCAAAAATACATATATG 3'	-18.98
CNGA1	3' GTGTGT-GGT-TGT-TATGTATATAT 5' : : 5' TATATATCTATATAGGCACATATATA 3'	-16.12
18h18	3' GTGTGTG-GT-TGT-TATGTATATAT 5' : : 5' AATATATATATATATACATATATA 3'	-16.63
33h20	3' GGTGTGTG-GT-TGT-TATGTATATAT 5' : : 5' TATATGTATATATATACATATATA 3'	-18.57
SETD6	3' GTGTGTG-GT-TGT-TATGTATATAT 5' : : 5' CATGTATACATATGTATACATATATG 3'	-15.76
Aoc2	3' GTGTG-TGGT-TGT-TATGTATATAT 5' : : 5' CAAACTACTATGCATTATACATATATG 3'	-16.62

Table 2: Continued

Genes	gga-mir-1657-A	kcal/mol
BRD1	3' GTGTGTGGTTG-TTAT-GTATATAT 5' : : 5' AA-AC-CCAAACAAAACCATATATA 3'	-14.71
LMF1	3' G-TGT-GTGG-T-TG-T-T-ATGTATATAT 5' : : 5' CTACAGTACTAAGACAATGTTACATATATA 3'	-18.96
8j5	3' GTGTGTGG-TT-GTTATGTATATAT 5' : : 5' CTTATGCCAAAGC-ATGCATATATA 3'	-14.07
LOC419498	3' GTGTGT-GGT-TGTTATGTATATAT 5' : : 5' CATGGATTTGTAC-ATACATATATA 3'	-14.85
VAMP2	3' GTGTGTG-GT-TGIT-A-T-GTATATAT 5' : : 5' CATATGCGTATAACAAGTAAGCATATATA 3'	-20.58
LOC416335	3' GTGTGTGGTGT-TATGTATATAT 5' : : 5' TA-A-ACCAGGGG-ACATATATG 3'	-14.66
8d24	3' GTG-TG-TGGTTGT-TATGTATATAT 5' : : 5' CACTCCTGCTGTTTACACACATATA 3'	-15.26
FGFBP1	3' GTGTGTGGTGT-T-A-TGTATATAT 5' : : 5' TACAGATT-CC-ATTTGCACATATG 3'	-18.37
TERF1	3' GTGTGT-G-GTTTGTATGTATATAT 5' : 5' CAGA-ACCTCAAAAA-ACACATATG 3'	-15.23
SLC25A29	3' GTGTGT-GG-TT-G-T-T-A-TGTATATAT 5' : : 5' CA-ACATCTGGAGTAAGAACCACACATATA 3'	-15.18
SLC25A29	3' GTGTGTGGTGT-TATGTATATAT 5' : : 5' CATGTAGCAG-GA-GCACATATA 3'	-17
LATS1	3' GTGTGT-GGT-TGTATGTATATAT 5' : : 5' TATATATCTATATAG-GCACATATA 3'	-19

Table 3: Distribution of genotypes of A>G located in the seed regions of gga-mir-1657 in 6 chicken breeds

Breed ^a	Genotypes frequency (%)			Allele frequency (%)		Chi-square (χ^2) test	p-values
	GG	GA	AA	G	A		
WG	89.66	10.34	0	96.55	3.45	0.0566	0.8121
WL	100.00	0.00	0	100.00	0.00	-	-
JB	93.33	6.67	0	96.67	3.33	0.0182	0.8927
LH	66.67	33.33	0	83.33	16.67	1.0653	0.3020
BF	90.00	10.00	0	94.83	5.17	0.0566	0.8120
LY	90.00	10.00	0	95.00	5.00	0.0545	0.8154

^aBreed codes as showed in Table 1

AA GG GG GG GG GG AA GG GG AG AG AG M

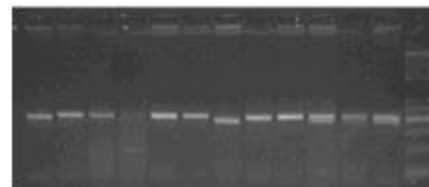


Fig. 4: PCR-RFLP patterns of the chicken mir-1657, M = Marker 50 bp ladder

expected heterozygosity were all <0.5 within the 6 chicken breeds; researchers concluded that genetic diversity was

Table 4: The Ewens-Watterson test for neutrality of A>G located in the seed regions of *gga-mir-1657* in 6 chicken breeds

Breed ^a	F statistical value of evolutionary power (observed F*)		95% confidence interval		
	Mean	SE	Lower bound	Upperbound	
WG	0.9019	0.7849	0.0270	0.5024	0.9661
WL	1.0000	-	-	-	-
JB	0.9334	0.7880	0.0267	0.5024	0.9661
LH	0.7222	0.7878	0.0270	0.5022	0.9672
BF	0.9019	0.7777	0.0278	0.5024	0.9661
LY	0.9050	0.7905	0.0267	0.5022	0.9672

^aBreed codes as showed in Table 1

Table 5: Genetic diversity of A>G located in the seed regions of *gga-mir-1657* in 6 chicken breeds

Breed ^a	Ne ^c	I ^d	F _{IS} ^e	Heterozygosity		
				Observed	Expected	Nei's ^b
WG	0.1034	0.0998	0.0981	1.1088	0.2036	-0.0545
WL	0.0000	0.0000	0.0000	1.0000	0.0000	-
JB	0.0690	0.0678	0.0666	1.0713	0.1500	-0.0357
LH	0.3333	0.2825	0.2778	1.3846	0.4506	-0.2000
BF	0.1000	0.0966	0.0950	1.1050	0.1985	-0.0526
LY	0.1000	0.0966	0.0950	1.1050	0.1985	0.4231
Average	0.1180	0.1113	0.1110	1.1249	0.2242	-0.1159

^aBreed codes as shown in Table 1, ^bNei's (1973) expected heterozygosity, ^cNe, effective number of alleles, ^dI, Shannon's information index, ^eFIS, Wright's fixation index

deficient at the detected site. The effective number of alleles and Shannon's information index represented the same trend as expected heterozygosity in the 6 chicken breeds. This might be the result of the directional artificial selection.

CONCLUSION

Findings presented in this study indicated that the *mir-1657* gene may be a galliformes-specific micro RNA. The rs14934924 SNP within the seed regions of *gga-mir-1657* gene maybe a functional sites which plays an important roles in the formation of some special phenotype including cancer susceptibility. Furthermore, the bioinformatics analysis also provides a basis for functional annotation of *gga-mir-1657* gene orthologs in other species.

ACKNOWLEDGEMENTS

This research was supported by National Science Foundation for Youths (No. 31001003) and Supported by Natural Science Foundation for Higher Education of Hebei Province (No. C2010129).

REFERENCES

Enright, A.J., B. John, U. Gaul, T. Tuschl, C. Sander and D.S. Marks, 2003. MicroRNA targets in *Drosophila*. *Genome Biol.*, 5: R1-R1.

Glazov, E.A., P.A. Cottee, W.C. Barris, R.J. Moore, B.P. Dalrymple and M.L. Tizard, 2008. A microRNA catalog of the developing chicken embryo identified by a deep sequencing approach. *Genome Res.*, 18: 957-964.

Liyang, G., C. Zhang and L. Du, 2009. Single nucleotide polymorphisms in chicken genomic pre-microRNA. *Biodiversity Sci.*, 17: 248-256.

Lutter, D., C. Marr, J. Krumsiek, E.W Lang and F.J. Theis, 2010. Intronic microRNAs support their host genes by mediating synergistic and antagonistic regulatory effects. *BMC Genom.*, 11: 224-224.

Mishra, P.J., P.J. Mishra, D. Banerjee and J.R. Bertino, 2008. MiRSNPs or MiR-polymorphisms, new players in microRNA mediated regulation of the cell: Introducing microRNA pharmacogenomics. *Cell Cycle*, 7: 853-858.

Ryan, B.M., A.I. Robles and C.C. Harris, 2010. Genetic variation in microRNA networks: The implications for cancer research. *Nat. Rev. Cancer*, 10: 389-402.

Siam, R., R. Harada, C. Cadieux, R. Battat, C. Vadnais and A. Nepveu, 2009. Transcriptional activation of the *Lats1* tumor suppressor gene in tumors of *CUX1* transgenic mice. *Mol. Cancer*, 8: 60-60.

Yeh, F.C., R.C. Yang and T. Boyle, 1999. Microsoft Windows Based Freeware for Population Genetic Analysis. POPGENE Version 1.31, University of Alberta, Edmonton.

Zuker, M., 2003. Mfold web server for nucleic acid folding and hybridization prediction. *Nucl. Acids Res.*, 31: 3406-3415.