INTERNATIONAL JOURNAL OF AGRICULTURE & BIOLOGY
ISSN Print: 1560–8530; ISSN Online: 1814–9596
19–0771/2019/22–6–1451–1458
DOI: 10.17957/IJAB/15.1221
http://www.fspublishers.org

Full Length Article

A Genome-Wide Survey, Model Selection, Phylogenetic Analysis and Protein-Protein Interactional Network Profile of the Metazoan PRKAG Genes from 22 Vertebrate Genomes

Wuyi Liu[1*], Huifang Lv[1] and Alireza Seidavi[2*]
[1]*Department of Biology Sciences, Biological and Food Engineering Faculty, Fuyang Normal University, China*
[2]*Department of Animal Science, Rasht Branch, Islamic Azad University, Rasht, Iran*
[*]For correspondence: lwui@163.com; lwycau@163.com; alirezaseidavi@iaurasht.ac.ir

Abstract

AMP-activated protein kinase (AMPK) is a cellular energy sensor which regulates organizational energy and phosphorylates diverse enzymes, playing key roles in determining different metabolic phenotypes and the energy metabolism equilibrium. This genome-wide survey identified and analyzed 266 unique AMPK gamma subunit (PRKAG) gene sequences from the recently updated genomes of human and 21 animals. After the model selection of phylogenetic analysis parameters, all the 266 protein sequences of these identified human and animal PRKAG genes were used to compute ML (Maximum Likelihood) phylogenetic tree under the evolutionary WAG+G and the protein-protein interactional network analysis using the STRING database. Subsequently, the phylogenetic analysis verified and found all these PRKAG gene sub-families from the amphibian, fish, avian, and mammal genomes formed large monotonous phylogenetic clusters except for three frog PRKAG sequences. Particularly, all the primate PRKAG genes formed the super-class of their sole phylogenetic clusters. Furthermore, the mapped network profile analysis revealed that these PRKAG genes were functionally assembled and/or enriched in the energy metabolism, immune responsive and/or inflammatory and myosin or myosin heavy chain interacting proteins related signaling pathways. In addition, most of these proteins are interrelated and formed a tight protein interactional network of energy metabolism. In conclusion, the observed data and information of protein-protein interaction will contribute to understanding and analyzing the possible protein functions and interacted pathways of animal PRKAG genes and/or other AMPK subunit genes. The metazoan PRKAG gene sequences and identified dataset may be beneficial for related developmental and physiological research topics. These findings may also be potentially applied to explore and resolve the critical issues of disease-induced loss that many medical researchers are taking into research.

Keywords: Genome-Wide Survey; Phylogenetic Analysis; Evolutionary Algorithm; Protein-Protein Interactional network

Introduction

AMP-activated protein kinase (AMPK) is implicated as a cellular energy sensor which regulates organizational energy and phosphorylates diverse enzymes, including protein effectors and metabolic enzymes. In many studies and reports, AMPK plays key roles in determining different metabolic phenotypes and the energy metabolism equilibrium of biological organisms as well as other ATP-binding cassette proteins (Li *et al*., 2019) or ADG related uncoupling proteins (Jiang *et al*., 2019). Particularly, it appears as the primary energy inductor and regulator of cellular energy homeostasis acting in the long term and/or the short term in eukaryotes and to ameliorate the metabolic distress by post-translational metabolic enzyme modifications (Hardie, 2007; Hardie *et al*., 2012; Ahmad *et al*., 2015; Garcia and Shaw, 2017; Neumann and Viollet, 2019; Rashtchizadeh *et al*., 2019). In fact, AMPK and its subunits are also regarded the potential sensor molecules concerned in human and animal diseases and biomedicine targets for treating metabolic decline in aging and metabolism-dependent disorders (Burkewitz *et al*., 2014; Ahmad *et al*., 2015; Li *et al*., 2015; Carling, 2017; Troncone *et al*., 2017; Willows *et al*., 2017; Neumann and Viollet, 2019; Rashtchizadeh *et al*., 2019).

In structure, AMPK is made up of a catalytic alpha (α) subunit, a non-catalytic regulatory beta (β) subunit, and a connector gamma (γ) subunit. AMPK is actually a particularly economic three-dimensional complex with its effective α and γ core subunits and the additional β C-terminal domain, in which the β C-terminal domain is shuttled in the other two subunits. The α subunit has the main site threonine 172 (T172) of phospho-dephospho regulation, whereas the β subunit is associated with glycogen metabolism in a phosphorylation-dependent manner via its carbohydrate-binding module (Ahmad *et al*., 2015;

To cite this paper: Liu, W., H. Lv and A. Seidavi, 2019. A genome-wide survey, model selection, phylogenetic analysis and protein-protein interactional network profile of the metazoan PRKAG genes from 22 vertebrate genomes. *Intl. J. Agric. Biol*., 22: 1451–1458

Garcia and Shaw, 2017; Neumann and Viollet, 2019; Rashtchizadeh *et al*., 2019). On the other hand, there are the cystathionine-β-synthase (CBS) motifs and other internal domains indwelling in the AMPK γ subunit (PRKAG) to form those two binding sites required for the enzymatic activities of specific allosteric activators of AMP (Burkewitz *et al*., 2014; Carling, 2017; Troncone *et al*., 2017; Neumann and Viollet, 2019; Rashtchizadeh *et al*., 2019).

At present, many studies in the human and animal genetics and development research suggested that these AMPK subunit genes were functionally important in the nutrient intake and physiological and pathological changes in the energy metabolism equilibrium of skeletal and cardiac muscle tissues of human and pig and chicken with both physiological and genetic evidences (Milan *et al*., 2000; Andersson, 2003; Zhao *et al*., 2006; Scheffler *et al*., 2011; Yin *et al*., 2012; Scheffler *et al*., 2016; Yang *et al*., 2016). Overall, PRKAG2 mutations lead to a decrease in the ability of AMP to bind AMPK, resulting in excessive storage of glycogen and causing disease (Gollob *et al*., 2001; Burwinkel *et al*., 2005; Carling, 2017; Rashtchizadeh *et al*., 2019). For example, in human genetics, previous studies have shown that the genetic mutated locus from Arg 302 to Gln 302 in the PRKAG2 gene usually rendered the human AMPK γ2 subunit to be a structurally inactive or deficient kinase, resulting in the pre-excitation syndrome (Gollob *et al*., 2001), *i.e*., Wolff-Parkinson-White syndrome. Furthermore, the fatal congenital glycogen hyperactivity was reported as mainly caused by the R531Q locus mutation of human PRKAG2 gene (Burwinkel *et al*., 2005). In the Hampshire pigs, Milan *et al*. (2000) found and reported a non-conservative single nucleotide polymorphism (*i.e*., a dominant *RN*-mutation) of R200Q locus of the swine PRKAG3 alleles (Milan *et al*., 2000; Andersson, 2003; Scheffler *et al*., 2011, 2016), which was identified as significantly associated with the excess glycogen content in pork. Later, many mutant substitutions like the V199I and AMPKγ3R200Q loci of the swine PRKAG3 alleles and mutant substitutions of the bovine PRKAG1 alleles were also identified and evaluated in different pig populations in recent years (Scheffler *et al*., 2011, 2016). Since then, similar mutant allele cases of PRKAG genes reported in animal muscle tissues were also characterized in the allele loci of chicken PRKAG3 gene (Zhao *et al*., 2006; Yin *et al*., 2012; Yang *et al*., 2016) and dairy cows' PRKAG1 (Ahmad *et al*., 2015). Therefore, the genetic mutations of many PRKAG alleles have been the hotspot subject of domestic animal researchers in the genomics era.

Previously, an initial study has been conducted to mainly identify avian PRKAG genes using the genome-wide BLAST searches (Liu, 2017). However, more and more sequences of PRKAG and related genes remain to be further explored and classified, whereas their functional interactional network and related signaling pathways remain to be characterized, especially these in domestic animal genomes. The present study conducted an updated genome-wide survey and the large-scale phylogenetic analysis to further analyze and characterize the human and animal PRKAG genes identified from 22 vertebrate genomes. Furthermore, the study also carried out the mapped interaction network profile analysis of PRKAG genes.

Materials and Methods

Data Retrieval and Sequence Alignment of Human and Animal PRKAG Genes from the Genomic Databases

The genome-wide survey was carried out using human PRKAG genes as input query sequence of BLAST searches. That is, with a few pair nucleotide and protein sequences of human PRKAG genes, we systematically performed previously used PSI-BLAST searches (Luo *et al*., 2019) to retrieve all the putative PRKAG genes against the latest released genomes of 22 vertebrates in the NCBI genome resources (URL: http://www.ncbi.nlm.nih.gov/genome/guide/build.shtml), *i.e*., *Acinonyx jubatus, Bos taurus, B. mutus , Bubalus bubalis, Canis lupus, C. lupus familiaris, Capra hircus, Gorilla gorilla, Homo sapiens, Mus musculus, Ovis aries, Pongo abelii, Rattus norvegicus, Sus scrofa, Panthera tigris, Gallus gallus, Danio rerio, Meleagris gallopavo, Anas platyrhynchos, Taeniopygia guttata, Xenopus tropicalis, X. laevis*. In total, we had retrieved more than 600 sequences of the nucleotide and protein sequences of the putative PRKAG genes (*i.e*., *PRKAG1*, *PRKAG2*, and *PRKAG3*) and their isoform proteins. All of the retrieved nucleotide and protein sequences of vertebrate PRKAG genes were subsequently used in the sequence alignments that were transformed into FASTA sequence files with the Windows software package of ClustalX version 2.0 (Larkin *et al*., 2007). Meanwhile, the incomplete sequences and highly divergent sequences were excluded from the further analysis, because the incomplete sequences and sequence indels and/or gaps those sequences of would result in uncertain alignments.

Model Selection of Phylogenetic Tree Parameters of Vertebrate PRKAG Genes

MEGA version 6.06 (Tamura *et al*., 2013) was used to make model selection analysis of phylogenetic tree parameters for nucleotide and amino acid sequences. By default, MEGA presents for each model the estimated values of Bayesian information criterion (BIC), AICc (Akaike Information Criterion, corrected), Maximum Likelihood ratio (lnL), the estimated values of shape parameter of the discrete Gamma distribution (+G) and other parameters (including invariant sites and tree branch lengths) as applicable (Nei and Kumar, 2000; Tamura *et al*., 2013). Whenever applicable, the software package MEGA shows the estimates of gamma

shape parameter and/or the estimated fraction of invariant sites (Nei and Kumar, 2000; Tamura *et al*., 2013), whereas the non-uniformity of evolutionary rates among sites are computed under a discrete distribution of Gamma (+G) in the establishing of evolutionary phylogenetic trees (Tamura *et al*., 2013). Depending on the model, the goodness-of-fit of each model is mainly measured by the Bayesian information criterion (BIC) and the corrected Akaike information criterion (AICc) in calculation (Tamura *et al*., 2013). The simulated model parameters of phylogenetic analysis are briefly shown in Table 1 and the detailed phylogenetic model parameters provided in Supplementary Table 2.

Phylogenetic Analyses of Vertebrate PRKAG Genes

The evolutionary fit of major substitution models of phylogenetic tree parameters were compared and selected for nucleotide and amino acid sequences of vertebrate PRKAG genes (see the model selection part or simulated phylogenetic model parameters provided in Supplementary Table 2 for details). The finally identified putative protein sequences of vertebrate PRKAG genes were also used to compute the ML phylogenetic trees under the optimized evolutionary substitution model measured by the dual criterions of BIC and AICc (Nei and Kumar, 2000; Tamura *et al*., 2013), respectively. Among the subsequent analyses, ML phylogenetic trees were computed and reconstructed with the subtree-pruning-regrafting search heuristic method in all the established trees (Tamura *et al*., 2013).

Protein Interactional Network Analysis using the STRING Database

The Search Tool for the Retrieval of Interacting Genes (STRING; URL: https://string-db.org/) is a pre-built database of abundant known experimental and predicted protein interactions (Szklarczyk *et al*., 2015, 2019). To investigate all the possible protein–protein interactions between the vertebrate PRKAG genes retrieved and identified from the updated genome-wide survey in 22 vertebrates, the STRING database was used for the creation of protein–protein interactional network (PIN) files as previously described (Szklarczyk *et al*., 2015, 2019). In practice, the STRING database version 11 online was used to evaluate and analyze the protein–protein interactions of vertebrate PRKAG1 and PRKAG2 and PRKAG3 protein sequences (Szklarczyk *et al*., 2019). The finally created and extracted PIN network files were explored and compared based on the PRKAG gene expression products (protein sequences) and their PIN network interactions and/or topological characteristics reported by more than 20 reports or studies of vertebrate PRKAG genes.

Results

Genome-wide Survey and Identification of Vertebrate PRKAG Genes

In total, more than 600 sequences were initially retrieved and screened from the NCBI genome resources (URL: http://www.ncbi.nlm.nih.gov/genome/guide/build.shtml). The nucleotide and protein sequences of vertebrate PRKAG genes were subsequently used and aligned by ClustalX 2.0 (Larkin *et al*., 2007). Those incomplete sequences and highly divergent sequences were excluded from the further analysis, since the incomplete sequences and sequence indels and/or gaps those sequences of would result in uncertain alignments. Furthermore, other removed records are predicted transcription factors or hypothetical proteins. After removing these incomplete sequences and uncertain and/or redundant sequence records, we then obtained 266 unique protein sequences and 266 unique nucleotide sequences of the putative PRKAG genes of human and animals for further analyses. It should be noted that all the 266 pairs of unique protein and nucleotide sequences retrieved from the genomes of 22 vertebrate species were carefully checked and identified and verified again. Finally, these 266 pairs of unique PRKAG protein and nucleotide sequences were further checked and examine by all the members of our group. These final used sequence dataset included a total of 266 unique protein sequences from the genomes of 22 different animal species (Supplementary Table 1), including 16 sequences from *A. jubatus* (Cheetah), 18 sequences from *B. taurus* (cattle, bovine), 5 sequences from *B. mutus* (wild yak), 18 sequences from *B. bubalis* (water buffalo), 15 sequences from *C. lupus* (gray wolf), 18 sequences from *C. lupus familiaris* (dog), 18 sequences from *C. hircus* (goat), 10 sequences from *G. gorilla* (western gorilla), 22 sequences from *H. sapiens* (human), 15 sequences from *M. musculus* (mouse), 13 sequences from *O. aries* (sheep), 8 sequences from *P. abelii* (Sumatran orangutan), 13 sequences from *R. norvegicus* (rat), 14 sequences from *S. scrofa* (pig), 4 sequences from *P. tigris* (tiger), 13 sequences from *G. gallus* (chicken), 13 sequences from *D. rerio* (zebrafish), 3 sequences from *M. gallopavo* (turkey), 11 sequences from *A. platyrhynchos* (duck), 3 sequences from *T. guttata* (zebra finch), 10 sequences from *X. tropicalis* (Western clawed frog) and 6 sequences from *X. laevis* (African clawed frog).

The corresponding information of these 266 pairs of unique protein and nucleotide sequences identified for further analyses of the putative PRKAG genes of human and animals are summarized in Supplementary Table 1.

Model Selection and Phylogenetic Analyses of Vertebrate PRKAG Genes

In the study, the analyzed ML phylogenetic trees were used to identify and elucidate the evolutionary and functional

Table 1: Model selection of phylogenetic analysis parameters of animal PPKAG protein sequences by the MEGA software package

Model	Parameter	BIC	AICc	lnL	Invariant	Gamma
WAG+G	180	9838.5107	8414.4932	-4025.6445	n/a	4.5237
WAG+G+I	181	9848.4397	8416.5290	-4025.6445	0.0000	4.5237
WAG	179	9866.4203	8450.2963	-4044.5638	n/a	n/a
JTT+G	180	9868.0000	8443.9825	-4040.3892	n/a	4.2164
WAG+I	180	9876.3491	8452.3317	-4044.5637	0.0000	n/a
JTT+G+I	181	9877.9290	8446.0183	-4040.3892	0.0000	4.2164
LG+G	180	9881.7111	8457.6936	-4047.2447	n/a	3.9697
LG+G+I	181	9891.6401	8459.7294	-4047.2447	0.0000	3.9697
JTT	179	9901.0844	8484.9604	-4061.8959	n/a	n/a
JTT+I	180	9911.2548	8487.2374	-4062.0166	0.0027	n/a
LG	179	9918.7117	8502.5877	-4070.7095	n/a	n/a
rtREV+G	180	9922.5346	8498.5171	-4067.6565	n/a	4.2164
LG+I	180	9928.6407	8504.6232	-4070.7095	0.0000	n/a
rtREV+G+I	181	9932.4636	8500.5529	-4067.6565	0.0000	4.2164
Dayhoff+G	180	9944.9559	8520.9385	-4078.8671	n/a	3.8518
Dayhoff+G+I	181	9954.9403	8523.0296	-4078.8948	0.0019	3.8907
rtREV	179	9956.4594	8540.3353	-4089.5834	n/a	n/a
rtREV+I	180	9967.0431	8543.0257	-4089.9107	0.0000	n/a
Dayhoff	179	9981.5271	8565.4031	-4102.1172	n/a	n/a
cpREV+G	180	9989.1781	8565.1606	-4100.9782	n/a	3.5901
Dayhoff+I	180	9990.8810	8566.8636	-4101.8297	0.0036	n/a
cpREV+G+I	181	9999.1071	8567.1964	-4100.9782	0.0000	3.5901
cpREV	179	10047.7906	8631.6666	-4135.2490	n/a	n/a
cpREV+I	180	10057.8358	8633.8183	-4135.3071	0.0000	n/a
WAG+G+F	199	10142.5227	8568.5678	-4083.3250	n/a	4.4785
WAG+G+I+F	200	10152.4517	8570.6074	-4083.3250	0.0000	4.4785
WAG+F	198	10170.2020	8604.1368	-4102.1291	n/a	n/a
LG+G+F	199	10170.9823	8597.0274	-4097.5548	n/a	3.9300
JTT+G+F	199	10172.5870	8598.6321	-4098.3571	n/a	4.1743
rtREV+G+F	199	10174.0040	8600.0491	-4099.0656	n/a	4.0503
WAG+I+F	199	10180.1310	8606.1761	-4102.1291	0.0000	n/a
LG+G+I+F	200	10180.9113	8599.0670	-4097.5548	0.0000	3.9300
rtREV+G+I+F	200	10183.9335	8602.0892	-4099.0658	0.0000	4.0503
JTT+G+I+F	200	10185.3393	8603.4950	-4099.7687	0.0000	4.1325
LG+F	198	10209.4468	8643.3815	-4121.7515	n/a	n/a
LG+I+F	199	10220.1686	8646.2136	-4122.1479	0.0000	n/a
Dayhoff+G+F	199	10230.5047	8656.5498	-4127.3159	n/a	3.8518
Dayhoff+G+I+F	200	10240.0382	8658.1938	-4127.1182	0.0038	3.8907
Dayhoff+F	198	10265.5408	8699.4756	-4149.7985	n/a	n/a
JTT+F	198	10270.5448	8704.4795	-4152.3005	n/a	n/a
Dayhoff+I+F	199	10274.4136	8700.4587	-4149.2704	0.0048	n/a
rtREV+F	198	10281.9007	8715.8354	-4157.9784	n/a	n/a
rtREV+I+F	199	10284.2329	8710.2780	-4154.1800	0.0000	n/a
cpREV+G+F	199	10294.8042	8720.8492	-4159.4657	n/a	3.5901
JTT+I+F	199	10304.9600	8731.0051	-4164.5436	0.0048	n/a
cpREV+G+I+F	200	10306.7401	8724.8958	-4160.4692	0.0000	3.5901
mtREV24+G+F	199	10335.5066	8761.5517	-4179.8169	n/a	3.5901
mtREV24+G+I+F	200	10345.4356	8763.5912	-4179.8169	0.0000	3.5901
cpREV+F	198	10362.8536	8796.7883	-4198.4549	n/a	n/a
cpREV+I+F	199	10376.2941	8802.3392	-4200.2106	0.0007	n/a
mtREV24+G	180	10395.6196	8971.6022	-4304.1990	n/a	3.5901
mtREV24+F	198	10396.1125	8830.0472	-4215.0844	n/a	n/a
mtREV24+I+F	199	10401.1985	8827.2436	-4212.6629	0.0000	n/a
mtREV24+G+I	181	10405.5487	8973.6380	-4304.1990	0.0000	3.5901
mtREV24+I	180	10468.9269	9044.9094	-4340.8526	0.0000	n/a
mtREV24	179	10610.5703	9194.4462	-4416.6388	n/a	n/a

Note: The model abbreviations or captions are interpreted as follows. GTR: General Time Reversible; JTT: Jones-Taylor-Thornton; rtREV: General Reverse Transcriptase; cpREV: General Reversible Chloroplast; mtREV24: General Reversible Mitochondrial; n/a: not available

divergence of these PRKAG genes in 22 different species. In the surveyed genomes of 22 vertebrate species, totally 266 pair unique putative PRKAG nucleotide and protein sequences were finally identified and used in the phylogenetic analysis (Supplementary Table 1). In the ML phylogenetic tree of the 266 unique putative protein sequences of vertebrate PRKAG genes, almost all the protein sequences were clustered into four large classes of amphibian, fish, avian, and mammal PRKAG genes accordingly (Fig. 1). However, a few exceptions do exist in the ML phylogenetic tree (Fig. 1). From all the three sub-families of PRKAG sequences (*i.e.*, *PRKAG1*, *PRKAG2*, and *PRKAG3*), the PRKAG sequences identified from the amphibian, fish, avian, and mammal genomes formed large

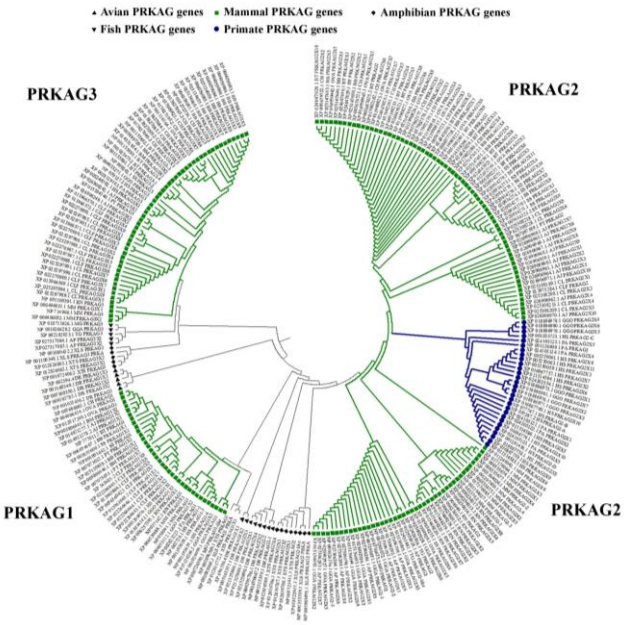

Fig. 1: ML phylogenetic tree of the 266 protein sequences of vertebrate PRKAG genes

Note: It is the circular view of ML phylogenetic tree of human and animal PRKAG protein sequences. The phylogenetic tree was computed and reconstructed with these 266 putative protein sequences of PRKAG genes identified in the genomes of 22 vertebrate species

monotonous phylogenetic clusters except for three frog PRKAG sequences (NP_001096506.1, NP_001085968.1, and XP_012811772.1) that were close to those of fish. Furthermore, all the primate PRKAG sequences formed a sole super-class of their own phylogenetic clusters in blue circles (Fig. 1). In brief, the ML phylogenetic analysis of putative vertebrate PRKAG protein sequences revealed that these PRKAG gene subfamilies (*PRKAG1*, *PRKAG2* and *PRKAG3*) had their own independent common ancestor genes (Fig. 1).

Protein Interactional Network Analysis of Human and Animal PRKAG Protein Sequences

The protein–protein interaction of human and animal PRKAG protein sequences were analyzed and displayed using the STRING database (URL: http://string-db.org/) of experimental and predicted protein interactions. The STRING database is able to detect and find the molecular enrichment and functional correlation in a close protein interaction and experimentally validated contacted network among the genes implied the importance of any experimental subject pathways (Fig. 2). The picture was the evidence network view of key hub proteins (*e.g*., PRKAG1, PRKAG2 and PRKAG3) and their relevant partners (Figure 2). These interacted proteins mainly included GDE1 (Glycerophosphodiester phosphodiesterase 1), TCF3 (Transcription factor E2-alpha, a basic helix-loop-helix transcription factor

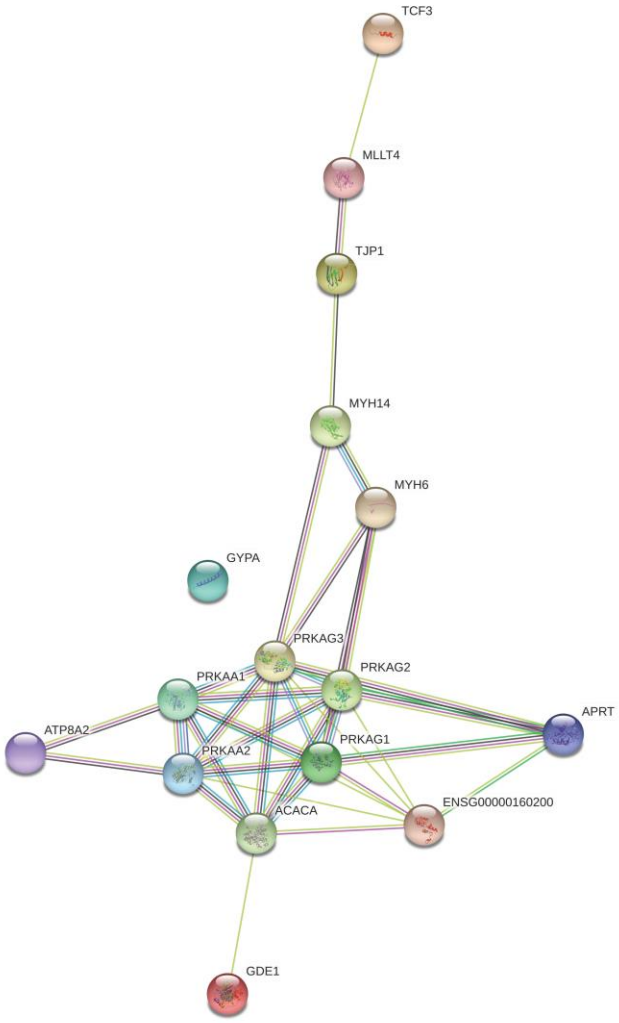

Fig. 2: The protein–protein interactional network (PIN) of human and animal PRKAG protein sequences by the STRING database

Note: It is the evidence view of protein–protein interactional network (PIN) built by the online bioinformatics tools of STRING database to probe the functional protein interactions and hub proteins of vertebrate PRKAG genes. Red line indicates the presence of fusion evidence, whereas green line represents neighborhood evidence. Blue line represents co-occurrence evidence, whereas purple line represents experimental evidence. Yellow line represents text mining evidence, whereas light blue line represents database evidence and a black line represents co-expression evidence. The average confidence scores for all of the protein–protein interactions is above 0.15 and those different line colors represent different types of protein functional associations (see the illustration shown the online website of the STRING database)

involved in the initiation of neuronal differentiation), TJP1 (Tight junction protein ZO-1 belonging to the MAGUK family whose N and C terminals may be involved in the cellular tight junction assembly), GYPA (Glycophorin-A, a receptor for influenza virus and a major intrinsic membrane protein critical for the function of SLC4A1), APRT (Adenine phosphoribosyltransferase catalyzing salvage reactions resulting in the formation of AMP), ATP8A2 (Phospholipid-transporting ATPase IB, a catalytic component of a P4-ATPase flippase complex), MLLT4 (Afadin, belonging to an adhesion system playing a

role in the organization of homotypic, interneuronal and heterotypic cell-cell adherens junctions), ENSG00000160200 (Cystathionine-beta-synthase), MYH6 (Myosin-6, belonging to the Myosin family), MYH14 (Myosin-14, belonging to the Myosin family), ACACA (Acetyl-CoA carboxylase 1), PRKAA1 (AMP-activated protein kinase catalytic subunit alpha-1, a catalytic subunit of AMPK), PRKAA2 (AMP-activated protein kinase catalytic subunit alpha-2, a catalytic subunit of AMPK), PRKAG1 (AMP-activated protein kinase subunit gamma-1, the AMP/ATP-binding subunit of AMPK), PRKAG2 (AMP-activated protein kinase subunit gamma-2, the AMP/ATP-binding subunit of AMPK), and PRKAG3 (AMP-activated protein kinase subunit gamma-3, the AMP/ATP-binding subunit of AMPK). Besides GYPA (Glycophorin A), most of these protein partners (encoded by corresponding genes) are interrelated and formed a tight protein interactional network of energy metabolism (Fig. 2). In particular, the main hub proteins of the protein–protein interactional network are PRKAG1, PRKAG2, PRKAG3, PRKAA1, PRKAA2, APRT, ATP8A2, ENSG00000160200, MYH6, MYH14 and ACACA, interacting with no less than 3 protein partners or enzymes. Interestingly, they also linked to many immune responsive and/or inflammatory factors and myosin proteins (Fig. 2), such as MYH6 (myosin 6) and MYH14 (myosin 14).

Discussion

In this study, we identified and characterized 266 unique PRKAG genes from the updated genomes of human and 21 animals. In the present study, accurate ML phylogenetic analysis was used to identify putative homologous and/or orthologous relationships of these PRKAG genes in 22 different species with the presupposed corresponding model selection of phylogenetic trees. The phylogenetic clusters of PRKAG gene family analyzed here were in agreement much with the physiological and genetics classification of previous studies and reports with a more diversity observed among different vertebrate species. As mentioned in the corresponding part of materials and methods, these 266 finally identified putative protein sequences of human and animal PRKAG gene were also used to compute the ML phylogenetic trees under the optimized evolutionary WAG+G model measured by the dual criterions of BIC and AICc as shown in Table 1 and Supplementary Table 2 (Nei and Kumar, 2000; Tamura *et al*., 2013), respectively.

In practice, the protein–protein interaction of human and animal PRKAG protein sequences were analyzed using the STRING database of experimental and predicted protein interactions. The mapped integrative profile of STRING protein network was built to probe into the functional protein interactions of vertebrate PRKAG genes (*e.g*., PRKAG1, PRKAG2 and PRKAG3) and their relevant partners. Interestingly, these interacted proteins were linked to many immune responsive and/or inflammatory factors and myosin proteins (Fig. 2), such as MYH6 (myosin 6) and MYH14 (myosin 14), whereas the previous reports indicated that PRKAG results in phosphorylation of the myosin heavy chains (Williams and Coluccio, 1995; Rosenberg and Ravid, 2006; Yan *et al*., 2018) and there was a cross talk between circadian rhythm and coronary heart disease identified by multiple correlation analysis (Yan *et al*., 2018). This phenomenon is in accord with previous studies that both the muscle and non-muscle myosin proteins are involved in the phosphorylation of ATP (Adenosine Triphosphate) and responses in energy metabolism and inflammatory (Williams and Coluccio, 1995; Musi *et al*., 2003; Rosenberg and Ravid, 2006; Yan *et al*. 2018; Vilchinskaya *et al*., 2018), which is evident to be associated with the function and activities of PRKAG genes. A recent review regarded the AMP-activated protein kinase genes (especially the *PRKAG* genes) as the key triggers for the disuse-induced skeletal muscle remodeling on some molecular targets (Vilchinskaya *et al*., 2018). In addition, there was a significant association of AMPK subunit gamma subunit (*PRKAG*) gene polymorphisms with growth, feed intake, and feed efficiency in meat-type chickens (Jin *et al*., 2016).

Meanwhile, the online resources of STRING database provide both these experimental data as well as the bioinformatics data and those predicted interaction information with confidence scores. The PIN profile of PRKAG kinases and their protein partners suggest that these proteins associated with short distances to each other in the functionally interacted network are more likely to share the common biological functions (Szklarczyk *et al*., 2015; Yan *et al*., 2018; Szklarczyk *et al*., 2019), whereas those interactive neighbors are more likely to have identical biological function than non-interactive ones (Szklarczyk *et al*., 2015; Yan *et al*., 2018; Szklarczyk *et al*., 2019). It is inferred that the query kinases and their interactive protein partners may form a complex to perform particular functions or involved in the same functional pathways. In brief, these STRING experimental and predicted protein interaction data and corresponding resulted protein interactional network will contribute to understanding and analyzing the possible protein functions and interacted pathways of the PRKAG genes and/or other AMPK subunit genes.

Conclusion

In the present study, phylogenetic analyses indicated that all the identified gene sequences belong to three sub-families of PRKAG genes forming four large super-classes of monotonous phylogenetic clusters (*i.e*., the amphibian, fish, avian, and mammal monotonous clusters). In particular, those PRKAG gene sequences identified from the amphibian, fish, avian and mammal genomes formed large monotonous phylogenetic clusters except for three frog

PRKAG sequences that were close to those of fish. Furthermore, the mapped network profile of PIN analysis revealed that the PRKAG genes were functionally assembled and/or enriched in energy metabolism, immune responsive and/or inflammatory, and myosin interacting proteins related signaling pathways. These observed data and information of protein-protein interaction will contribute to understanding and analyzing the possible protein functions and interacted pathways of vertebrate PRKAG genes and/or other AMPK subunit genes. The resulted findings will also be potentially applied to explore and resolve the critical issues of disease-induced loss that many medical researchers are taking into research.

Acknowledgments

This work was supported by the Major Projects of Horizontal Cooperation between Fuyang Municipal Government and Fuyang Normal University (No. XDHX201725), Anhui Provincial Educational Commission Natural Science Foundation (No. KJ2019ZD36) and Anhui Provincial Scientific and Technological Project (Grant Number 201806a02020028), the Major Construction Discipline Project in Anhui Province, and the Biology First-Class Construction Discipline Project of Fuyang Normal University (No. 19XJS0710).

References

Ahmad, M., Z. Amir, A. Hamid-Reza, A.R.V.H. Assad and E. Shahin, 2015. A SNP in the 3'-untranslated region of ampkγ1 may associate with serum ketone body and milk production of holstein dairy cows. *Gene*, 574: 48–52

Andersson, L., 2003. Identification and characterization of AMPK gamma 3 mutations in the pig. *Biochem. Soc. Trans.*, 31: 232–235

Burkewitz, K., Y. Zhang and W. Mair, 2014. AMPK at the Nexus of Energetics and Aging. *Cell Metabol.*, 20: 10–25

Burwinkel, B., J.W. Scott, B. Christoph, F.K.H.V. Landeghem, G.F. Cox, C.J. Wilson, D.G. Hardie and M.W. Kilimann, 2005. Fatal Congenital Heart Glycogenosis Caused by a Recurrent Activating R531Q Mutation in the γ2-Subunit of AMP-Activated Protein Kinase (PRKAG2), Not by Phosphorylase Kinase Deficiency. *Amer. J. Hum. Genet.*, 76: 1034–1049

Carling, D., 2017. AMPK signalling in health and disease. *Curr. Opin. Cell. Biol.*, 45: 31–37

Garcia, D. and R.J. Shaw, 2017. AMPK: mechanisms of cellular energy sensing and restoration of metabolic balance. *Mol. Cell*, 66: 789–800

Gollob, M.H., M.S. Green, A.S.L. Tang, T. Gollob, A. Karibe, A.S. Ali Hassan, F. Ahmad, R. Lozado, G. Shah, L. Fananapazir, L.L. Bachinski, T. Tapscott, O. Gonzales, D. Begley, S. Mohiddin and R. Roberts, 2001. Identification of a gene responsible for familial Wolff-Parkinson-White syndrome. *New Engl. J. Med.*, 344: 1823–1831

Hardie, D.G., 2007. Amp-activated/snf1 protein kinases: conserved guardians of cellular energy. *Nat. Rev. Mol. Cell. Biol.*, 8: 774–785

Hardie, D.G., F.A. Ross and S.A. Hawley, 2012. AMPK: a nutrient and energy sensor that maintains energy homeostasis. *Nat. Rev. Mol. Cell. Biol.*, 13: 251–262

Jiang, Q.F., L. Hong, Y. Feng, L. Chen, J. Zhang, L. Wu and Y.L. Gu, 2019. Correlation between uncoupling protein3 (UCP3) gene polymorphism and growth traits of hybrid simmental. *Intl. J. Agric. Biol.*, 21: 61–66

Jin, S., E.M.E. Moujahid, Z. Duan, J. Zheng, L. Qu, G. Xu, N. Yang and S. Chen, 2016. Association of AMPK subunit gene polymorphisms with growth, feed intake, and feed efficiency in meat-type chickens. *Poult. Sci.*, 95: 1492–1497

Larkin, M.A., G. Blackshields, N.P. Brown, R. Chenna, P.A. McGettigan, H. McWilliam, F. Valentin, I.M. Wallace, A. Wilm, R. Lopez, J. Thompson, T.J. Gibson and D.G. Higgins, 2007. Clustal W and Clustal X version 2.0. *Bioinformatics*, 23: 2947–2948

Li, N.N., Z.C. Xiao, M.W. Guan, S.T. Li, X.H. Zhang and F. Luo, 2019. Genome-Wide Analysis of the ATP-Binding Cassette A (ABCA) Gene Family and their Expression Profiles in High and Low Oil Cultivars of *Brassica napus*. *Intl. J. Agric. Biol.*, 21: 597–605

Li, X., L. Wang, X.E. Zhou, J. Ke, P.W.D. Waal, X. Gu, M.H. Tan, D. Wang, D. Wu, H.E. Xu and K. Melcher, 2015. Structural basis of AMPK regulation by adenine nucleotides and glycogen. *Cell Res.*, 25: 50–66

Liu, W.Y., 2017. Genome-wide Identification and Analysis of Human and Avian 5'-AMP-Activated Protein Kinase Gamma Subunit Genes. *Biotechnology*, 16: 145–154

Luo, B., J. Wang, H. Guo, Z.G. Yang and H.Y. Sun, 2019. Identification of Rice Lipid a Biosynthetic Genes and Expression Profile for Potential Roles in Various Physiological Processes. *Intl. J. Agric. Biol.*, 21: 325–332

Milan, D., J.T. Jeon, C. Looft, V. Amarger, A. Robic, M. Thelander, C. Rogel-Gaillard, S. Paul, N. Iannuccelli, L. Rask, H. Ronne, K. Lundström, N. Reinsch, J. Gellin, E. Kalm, P.L. Roy, P. Chardon and L. Andersson, 2000. A mutation in PRKAG3 associated with excess glycogen content in pig skeletal muscle. *Science*, 288: 1248–1251

Musi, N., H. Yu and L.J. Goodyear, 2003. AMP-activated protein kinase regulation and action in skeletal muscle during exercise. *Biochem. Soc. Trans.*, 31: 191–195

Nei, M. and S. Kumar, 2000. *Molecular Evolution and Phylogenetics*. Oxford University Press, Book, New York, USA

Neumann, D. and B. Viollet, 2019. AMP-Activated Protein Kinase Signalling. *Intl. J. Mol. Sci.*, 20: 766

Rashtchizadeh, N., H. Argani, H.A. Ghorbani, D. Sanajou, V. Hosseini, S. Dastmalchi and S.A.S. Nazari, 2019. AMPK: a promising molecular target for combating cisplatin toxicities. *Biochem. Pharmacol.*, 163: 94–100

Rosenberg, M. and S. Ravid, 2006. Protein kinase Cgamma regulates myosin IIB phosphorylation, cellular localization, and filament assembly. *Mol. Biol. Cell*, 17: 1364–1374

Scheffler, T.L., S. Park, P.J. Roach and D.E. Gerrard, 2016. Gain of function AMP-activated protein kinase γ3 mutation, AMPKγ3R200Q. in pig muscle increases glycogen storage regardless of AMPK activation. *Physiol. Rep.*, 4: 1–9

Scheffler, T.L., S. Park and D.E. Gerrard, 2011. Lessons to learn about postmortem metabolism using the (AMPKγ3R200Q) mutation in the pig. *Meat Sci.*, 89: 244–250

Szklarczyk, D., A.L. Gable, D. Lyon, A. Junge, S. Wyder, J. Huerta-Cepas, M. Simonovic, N.T. Doncheva, J.H. Morris, P. Bork, L.J. Jensen and C.V. Mering, 2019. STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucl. Acids Res.*, 47: 607–613

Szklarczyk, D., A. Franceschini, S. Wyder, K. Forslund, D. Heller, J. Huerta-Cepas, M. Simonovic, A. Roth, A. Santos, K.P. Tsafou, M. Kuhn, P. Bork, L.J. Jensen and C.V. Mering, 2015. STRING v10: protein-protein interaction networks, integrated over the tree of life. *Nucl. Acids Res.*, 43: 447–452

Tamura, K., G. Stecher, D. Peterson, A. Filipski and S. Kumar, 2013. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Mol. Biol. Evol.*, 30: 2725–2729

Troncone, M., S.M. Cargnelli, L.A. Villani, N. Isfahanian, L.A. Broadfield, L. Zychla, J. Wright, G. Pond, G.R. Steinberg and T. Tsakiridis, 2017. Targeting metabolism and AMP-activated kinase with metformin to sensitize non-small cell lung cancer (NSCLC) to cytotoxic therapy; translational biology and rationale for current clinical trials. *Oncotarget*, 8: 57733–57754

Vilchinskaya, N.A., I.I. Krivoi and B.S. Shenkman, 2018. AMP-Activated Protein Kinase as a Key Trigger for the Disuse-Induced Skeletal Muscle Remodeling. *Intl. J. Mol. Sci*., 19: 1–19

Williams, R. and L.M. Coluccio, 1995. Phosphorylation of myosin-I from rat liver by protein kinase C reduces calmodulin binding. *Biochem. Biophys. Res. Commun.*, 216: 90–102

Willows, R., M.J. Sanders, B. Xiao, B.R. Patel, S.R. Martin, J. Read, J.R. Wilson, J. Hubbard, S.J. Gamblin and D. Carling, 2017. Phosphorylation of AMPK by upstream kinases is required for activity in mammalian cells. *Biochem. J.*, 474: 3059–3073

Yan, X., Y. Huang and J. Wu, 2018. Identify cross talk between circadian rhythm and coronary heart disease by multiple correlation analysis. *J. Comput. Biol.*, 25: 1312–1327

Yang, Y., D. Xiong, L. Yao and C. Zhao, 2016. An SNP in exon 11 of chicken 5'-AMP-activated protein kinase gamma 3 subunit gene was associated with meat water holding capacity. *Anim. Biotechnol*., 27: 13–16

Yin, H.D., S.Y. Chen, Y. Wang, X.L. Zhao, Q. Zhu and Y.P. Liu, 2012. Association of Protein Kinase Adenosine Monophosphate-activated γ3-subunit (PRKAG3) Gene Polymorphisms with Carcass Traits in Chinese Meat-type Chickens. *J. Poult. Sci*, 49: 254–259

Zhao, C., C. Wang, X. Deng, Y. Gao and C. Wu, 2006. Identification of single-nucleotide polymorphisms in 5' end and exons of the PRKAG3 gene in Hubbard White broiler, Leghorn layer, and three Chinese indigenous chicken breeds. *J. Anim. Breed. Genet*., 123: 349–352

[Received 07 May 2019; Accepted 29 Jul 2019; Published (online) 22 Dec 2019]