Genes & Genomics
https://doi.org/10.1007/s13258-019-00911-w

Online ISSN 2092-9293
Print ISSN 1976-9571

RESEARCH ARTICLE

Molecular evidence for adaptive evolution of olfactory-related genes in cervids

Hengxing Ba[2] · Tao Qin[2] · Zexi Cai[3] · Wenyuan Liu[2] · Chunyi Li[1,2]

Received: 27 June 2019 / Accepted: 20 December 2019

Abstract

Background Cervids have evolved very successful means for survival and thriving to adapt to various climates and environments. One of these successful means might be the effective and efficient way of communication. To support this notion, cervids are well equipped with a variety of skin glands that distribute in different body regions. However, studies relevant to adaptive evolution in cervids, particularly on olfactory reception at the molecular level, have thus far not been reported.

Objective To provide valuable insights into molecular evidence for the adaptive evolution of olfactory-related gene in cervids.

Methods Based on recently sequenced genomes of cervids and closely-related-species, we performed comparative genomic analysis at genome level using bioinformatics tools.

Results Tree topology strongly supported that Bovidae was the sister group of Moschidae and both formed a branch that was then clustered with Cervidae. Expansion of heavy chain genes of the dynein family and 51 rapidly evolving genes could be associated with adaptation of cilia that serve as sensory organelles and act as cellular antennae. Based on the branch-site model test along the deer branch spanning 7–21 mammalian species, 14 deer olfactory receptor genes were found to be undergoing positive selection pressure and 89 positive selection sites (probability > 60%) had amino acid substitutions unique to deer.

Conclusion This study, for the first time, provides significant molecular evidence for adaption of olfactory-related genes of cervids according to their olfactory behavior.

Keywords Cervidae · Comparative genomics · Adaptation · Olfactory · Positive selection gene · Gene family

Introduction

Some functional classes of genes repeatedly show signs of accelerated adaptive evolution in different animal groups, including gene gain, loss or mutation, which have resulted in alteration of immune defense, reproduction, cell signaling, stress response, metabolism, and chemoreception, etc. (Schrider and Hahn 2010). Cervidae constitute the second most diverse family of artiodactyla (four subfamilies, 16 genera and about 44 species) after Bovidae, and live in the most vast of biomes (from tundra to the tropical rainforest) (Groves 2007; Grubb 1993). Being able to achieve such remarkable status, cervids must have evolved very successful means for survival and thriving. One of these successful means might be the effective and efficient way of communication: to escape from predators, to mark out territory and attract opposite sexes etc. As the forest dwellers, most of deer species live in thick bushes, obviously vision would not be a good choice for the resolution.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s13258-019-00911-w) contains supplementary material, which is available to authorized users.

✉ Hengxing Ba
bahengxing@caas.cn

✉ Chunyi Li
Lichunyi1959@163.com

[1] Changchun Sci-Tech University, Changchun 130600, China

[2] State Key Laboratory for Molecular Biology of Special Wild Economic Animals, Institute of Special Wild Economic Animals and Plants, Chinese Academy of Agricultural Sciences, Changchun 130112, China

[3] Department of Molecular Biology and Genetics, Center for Quantitative Genetics and Genomics, Aarhus University, 8830 Tjele, Denmark

Published online: 04 January 2020

Whereas, secreting and sensing different types of odorant molecules could logically be evolved and utilized to fulfil the purpose. To support this notion, cervids are well equipped with a variety of skin glands that distribute in different body regions, including interdigital glands, forehead glands, tarsal glands, preorbital glands, caudal glands, nasal glands and others (Adams and Johnson 2010; Meyer et al. 2008; Wood 2003).

Odorant molecules in skin-gland secretions in the cervids provide multiple signaling messages to conspecific or members of other related species (Mykytowycz and Goodrich 1974; Ralls 1971; Thiessen and Rice 1976). In black-tailed deer, for instance, they mark their home ranges with secretions from the forehead glands, and when fleeing, leave on the ground an alarm scent from the interdigital glands (Müller-Schwarze et al. 1984). Based on these previous studies, we propose that odorant receptor genes in cervids are likely to be undergoing coevolution with the purpose of obtaining a complete feedback of these scent markings.

However, studies relevant to adaptive evolution in cervids, particularly on olfactory reception at the molecular level, have thus far not been reported. Recently, the genomes of several deer species, i.e. white-tailed deer (*Odocoileus virginianus*, WT deer), reindeer (*Rangifer tarandus*; Li et al. 2017), Milu (*Elaphurus davidianus*; Zhang et al. 2018) and red deer (*Cervus elaphus*; Bana et al. 2018), as well as closely related musk deer (*Moschus Bweezovskii*; Fan et al. 2018), have been sequenced. The acquired sequence data have provided genomic information to a wide variety of morphological and physiological studies. Here, we took advantage of these publically available whole genome sequences to provide valuable insights into molecular evidence for the adaptive evolution of olfactory-related gene in cervids.

Results

Phylogenetic relationship at genomic level

We compiled 5293 single-copy orthologous genes from 15 mammal species (WT deer, Milu, reindeer, musk deer, cattle, yak, chiru, goat, sheep, dolphin, whale, pig, camel, alpaca and horse). A phylogenetic tree was constructed from these concatenated ortholog alignments using the maximum likelihood method. Tree topology strongly supported that Bovidae was the sister group of Moschidae and both formed a branch that was then clustered with Cervidae (100% bootstrap supports) (Fig. 1a). Our results clearly demonstrated systematic phylogenetic relationship between these three families.

Gene families

Analysis of changes in gene families revealed 69 expanded and 2419 contracted gene families in the cervid genomes (Fig. 1a). A significant higher frequency of gene-contraction than gene-expansion events is noted previously (Olson 1999). We focused on the expansion of the dynein gene family ($p < 0.05$) that accounted for 23, 22 and 28 orthologous copies in the WT deer, Milu and reindeer respectively (Fig. 1b). These orthologs included 13 axonemal heavy chain proteins (DNAH1, DNAH2, DNAH3, DNAH5, DNAH6, DNAH7, DNAH8, DNAH9, DNAH10, DNAH11, DNAH12, DNAH14 and DNAH17) and two cytoplasmic heavy chain proteins (DYNC1H1 and DYNC2H1). In view of the above evidence, an unrooted neighbor-joining tree of 262 dynein orthologs from the 15 mammalian species was constructed, and visualized (Fig. 1c) by using the Evolview tool (He et al. 2016). Our results showed that the orthologs of dynein genes were closely grouped in each cervid species, though some may appear to be pseudogenes because of short sequences. These observations further indicate that the duplication of dynein genes would have occurred since the common cervid ancestor evolved.

Rapidly evolving genes

Pairwise dN/dS ratios (the rate of non-synonymous substitutions to the rate of synonymous substitutions) were used to infer whether deer protein-coding genes were accumulated during rapid selection pressure in comparison between the three-closely-related families, i.e. Cervidae, Moschidaea and Bovidae, which, we believe, was well resolved in this study. As accuracy of the dN/dS ratio depends on completeness of coding sequences, the WT deer genome was chosen as a representative of Cervidae. We calculated dN/dS ratios for 10,772 single-copy orthologous gene pairs among WT deer, musk deer and cattle. A total of 2780 genes exhibited elevation of the difference values (D values) of dN/dS in both WT deer/cattle and WT deer/musk deer as compared to musk deer/cattle (Table S2), implying that these genes could have been subjected to accelerated evolution in the WT deer genome. These genes were selected to perform the analysis of gene ontology (GO) biological process enrichment. Interestingly, among the top gene ontology categories, we found that cilium-related biological processes were highly enriched, such as "cilium organization" (adjusted $p = 5.7 \times 10^{-4}$), "cilium morphogenesis" (adjusted $p = 0.002$), "cilium assembly" (adjusted $p = 0.001$) and "protein localization to cilium" (adjusted $p = 0.09$) (Fig. 2a, Table S3). A total of 51 rapidly evolved genes, colored by red dots in Fig. 2b, were involved in these cilium-related biological processes. A wilcoxon test indicates that there

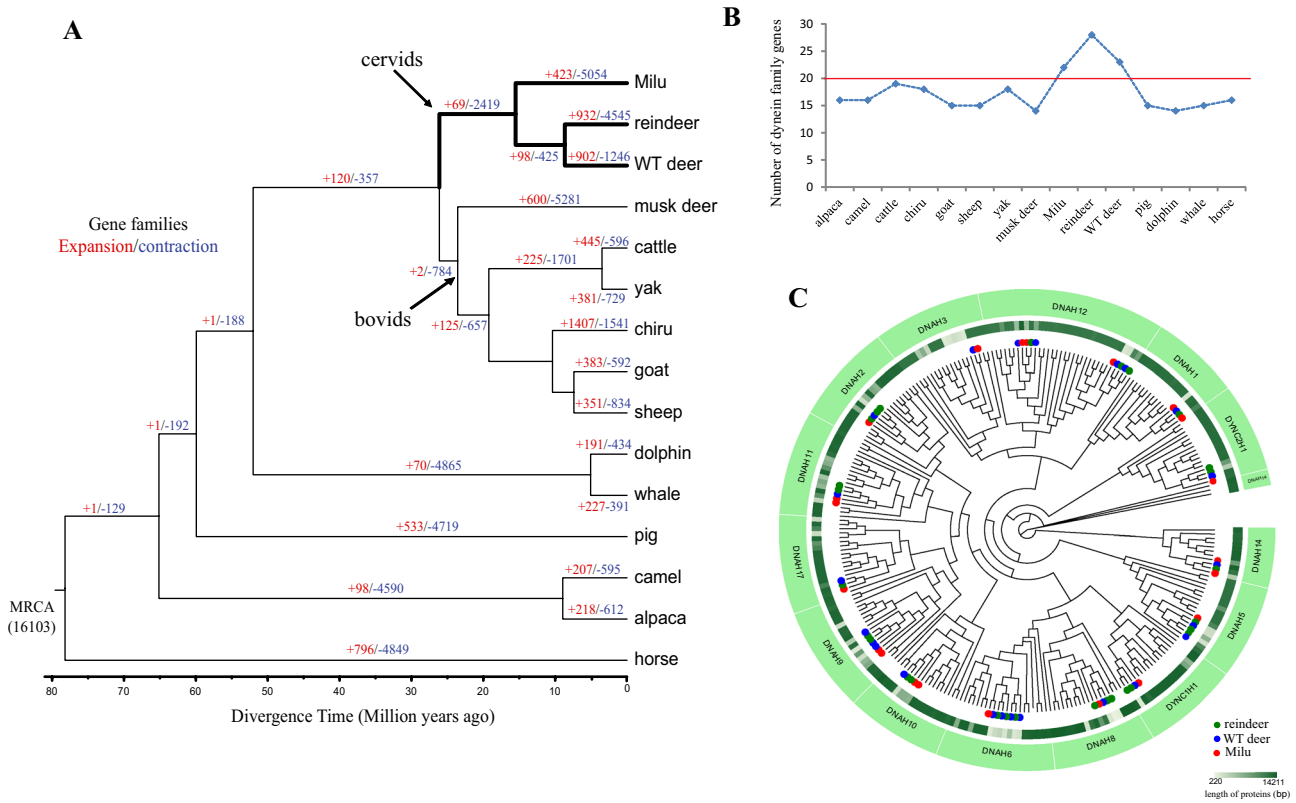

Fig. 1 Adaptation of gene families in cervid genomes. **a** Phylogenetic tree and divergence times estimated for the cervids and their relatives. Numbers associated with each branch designate for the numbers of gene families that have expanded (red) or contracted (blue) since the split from the common ancestor. **b** The number of orthologous genes within dynein family of closely related 15 mammalian species including WT deer, reindeer and Milu. **c** An unrooted neighbor-joining tree of dynein family in 15 mammalian species showing expanded dynein in the cervids (highlighted by colour dots) (colour figure online)

were significant differences in average D-value of dN/dS (both WT deer/cattle and WT deer/musk deer as compared to musk deer/cattle) between 51 cilium-related genes and 2780 background genes ($p = 0.043$) (Fig. 2c).

Positive selected olfactory-related genes

Branch-site model test was used to detect positive selected genes (PSGs) among the olfactory-related genes along WT deer branch ('foreground branch') spanning 7–21 mammalian species at the genome-scale, 11,160 orthologous genes in total. Through this analysis, 110 potential PSGs (FDR < 0.1) were identified in the WT deer genome, 14 out of the 110 genes (12.7%) belonged to olfactory receptor genes that were enriched to olfactory transduction pathway (KEGG, adjusted $p = 1.6 \times 10^{-2}$). Within these genes, 89 positive selection sites (probability > 60%) had amino acid substitutions unique to WT deer with frequency of average of six sites of each receptor (Table 1). Though these receptor genes and pseudogenes thereof vary enormously in mammalian genomes, we found our results were reliable based on manual checking.

Discussion

Our results showed that the cilium-related biological processes were highly enriched by the genes with the elevation of the D values of dN/dS radio. In vertebrates, there are two types of cilia: motile cilia and non-motile, or primary cilia. A substantial portion of primary cilia locates in dendritic knob of the olfactory neuron, and it typically serves as sensory organelles and acts as cellular antennae, providing sensation (e.g., odorant molecules) of extracellular environment (Adams et al. 2008). In our results, 51 potential rapidly evolving genes (e.g. ARL6) are involved in the process of ciliogenesis are likely to be correlated with the speedy adaptation of the olfactory nervous system in cervids. The motile cilia are usually present on a cell's surface in large numbers and beat in coordinated waves

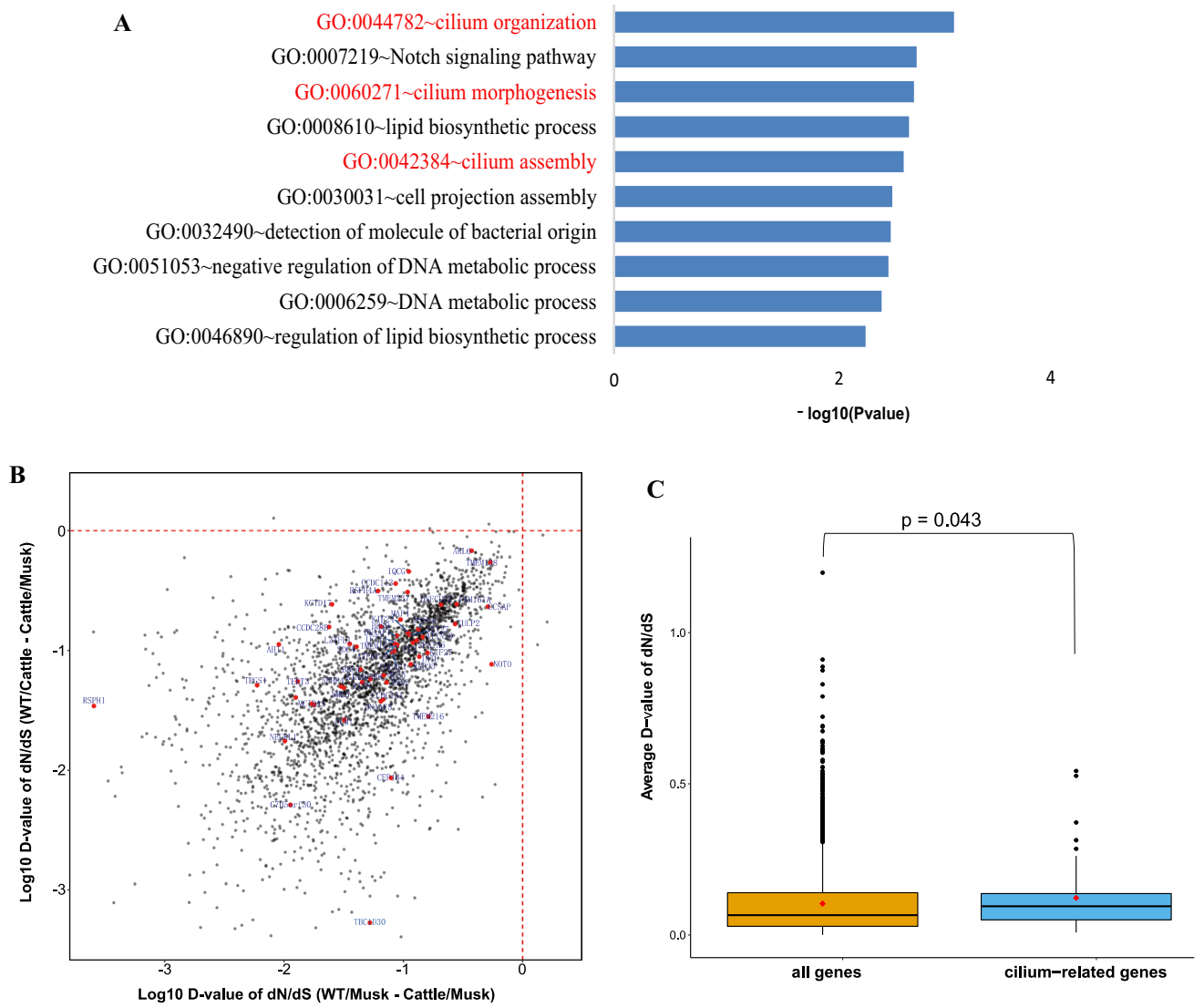

Fig. 2 Rapidly evolving genes on WT deer genome with the elevation of the dN/dS radio in both WT deer/cattle and WT deer/musk deer as compared with musk deer/cattle. **a** Top 10 DAVID GO biological process terms enriched by 2780 rapidly evolving genes. **b** Plot of both different values (D values) of dN/dS radio. Among these 2780 genes, 51 cilium-related genes were showed by red dots and labeled by blue fonts. **c** Boxplot of the D value of both 51 cilium-related genes and 2780 background genes. A wilcoxon test was performed between them (colour figure online)

(Lewin et al. 2007). The functioning of motile cilia is fully dependent on the presence of ciliary dynein. The latter is a family of cytoskeletal motor proteins, which convert the chemical energy stored in ATP to mechanical work to drive the beat of cilia. Ciliary dynein, also called axonemal dynein, is a member of dynein family, and comes in multiple forms consisting of heavy, light and intermediate chains (King and Schroer 2000). Regulation of the axonemal dynein activity is critical for cilia waveform (King 2010), which is important for the mediation of fluid flow to the odorant molecules to effectively reach ORs (Adams et al. 2008). In our results, expansions of the dynein heavy chains (e.g. DNAH6) further exhibit molecular evidence of adaptive evolution for physiological architecture of olfactory system in cervids. Here, the degree of expansion of the dynein gene family could be seriously underestimated in the present study because the used deer draft-genome-assemblies have sequence fragmentations and insufficient annotations.

Our results may also provide molecular evidence to support the hypothesis that ORs must be highly adaptive to selection and can distinguish great diversity of odorant sensing from skin glands. Moreover, a high percentage of the amino acid substitutions (36 out of the 89, 41%) between hydrophilic and hydrophobic (Table 1) is likely to have a degree of correlation between amino acid change

Table 1 Positive selection genes in cervids branch spanning mammals based on branch-site model test

Symbol (*Bos taurus*)	Positively selected sites with unique substitution (Probability > 60%)	Number of species	FDR	Description
LOC504567	I163V;V218A;I328T;S346K	7	0.06	Olfactory receptor 8S1
LOC507882	Y105C;S112A;V161L;H280R	18	0.08	Olfactory receptor 51F1-like
LOC513914	Q173R;E266K	13	0.08	Olfactory receptor 8B12
LOC515790	S97I;T138A;L148V;G151W;V257L;Q298H;	9	0.07	Olfactory receptor 4C46
LOC613418	V20M;V55L;L78F;V230G;	14	0.07	Olfactory receptor 12
LOC782191	K133Q;L231C;G260A;F279Y	12	0.06	Olfactory receptor 143
LOC782645	V288I;K302M	16	0.06	Olfactory receptor 10T2
LOC783845	I53V;P64C;A73S;A108G;N133T;S139V;F149T;S153T;A160C	10	0.00	Olfactory receptor 11H4
LOC788242	I12S;G15V;T73V;S74F;M80L;R121Q;L206V;H243Y;V275I	15	0.02	Olfactory receptor, family 6, subfamily C, member 4-like
LOC788438	L32S;D55N;S77F;T78S;K83R;A108P;M139T;C144R;S159F;K189T;S252G;Y281C;T282A	12	0.09	Olfactory receptor 12
OR10P1	T63A;M84V;Y85H;S92P;M144T;Y146C;C153F;Y158C;D206N;R212T;S221N;M234I;A253T;F264L;T280M;P288C;R298C;P309L	17	0.07	Olfactory receptor, family 10, subfamily P, member 1
OR52B6	Q39R;S46A;L177Q;L246I	15	0.09	Olfactory receptor, family 52, subfamily B, member 6
OR56A3	T220A;K244R;A245S;	15	0.04	Olfactory receptor, family 56, subfamily A, member 3
OR52E6	V118L;N130D;H173S;	9	0.02	Olfactory receptor, family 52, subfamily E, member 6

The underlined sites represent the amino acid substitutions between hydrophilic and hydrophobic

and the chemical composition of odorant molecules derived from the skin glands of cervids.

In conclusion, we believe our work has provided the evidence of adaptive evolution of olfactory-related genes in cervids including expansion of axonemal dynein genes for movement of motile cilia. The rapidly evolving genes involved in the process of ciliogenesis are also likely to be correlated with the speedy adaptation of the olfactory nervous system. The positive selection sites in OR genes could substantially increase sensitivity for chemoreception. Shortly before the completion of this manuscript, Ruminant genome sequencing project published 51 ruminant genomes (Chen et al. 2019), and therefore, more deer genomes should be added to investigate molecular evidence for adaptive evolution of olfactory-related genes in cervids in the future.

Materials and methods

Available genome sequences

Genomes from three deer species (white-tailed deer (WT deer), Milu and reindeer) and 20 other genomes from mammalian species (retrieved from the current Entrez Genome Project at NCBI or the genomic dataset at GigaDB) were used in this study (Table S1). Since the current red deer genome is only assembled to 1.96 Gbp in total, this genome was not included in this study.

Phylogenetic analysis

Gene families were constructed using closely related 15 mammalian species by OrthoMCL v2.0.9 (Li et al. 2003). This construction generated 5293 single-copy orthologs, which were aligned using MUSCLE v3.8 (Edgar 2004). Subsequently, 4D-sites (4-fold degenerated sites) of these alignments were extracted and concatenated to construct a maximum likelihood phylogenetic tree by RAxML v8.2.12 (Stamatakis 2014) with a GTR + G + I model and 100 bootstraps repeats. The divergence time for the analyzed taxa was estimated with MCMCtree, implemented in PAML v4.9 (Yang 2007), calibrated with published timings of the divergence of the reference species in TimeTree (http://www.timetree.org/, January 2018) (Hedges et al. 2015; Kumar et al. 2017). Finally, expansion and contraction of gene families was detected in these 15 mammalian species using CAFE v3.1 (Han et al. 2013) with P value threshold 0.05, number of random 10,000, and search for the λ value. Change in gene family was analyzed by using a CAFE report analysis script (https://hahnlab.github.io/CAFE/).

Analysis of rapid evolving genes

The ratio of dN/dS of single-copy orthologous pairs among WT deer, musk deer and cattle were calculated using KaKs_Calculator (Zhang et al. 2006), implemented in ParaAT v2.0 (Zhang et al. 2012), a tool that paralleled constructs protein-coding sequence alignments and calculates the ratio of dN/dS for a large number of orthologs. DAVID v6.8 online bioinformatics tool (Huang et al. 2009) was used to perform GO biological process annotation analysis.

Analysis of positive selected genes

To test for signatures of positive selection acting on the cervid branch of a phylogenetic tree, we used PosiGene v0.1 (Sahm et al. 2017), an automated parallel pipeline for genome-wide detection by comparing the likelihood scores of branch-site model using CODEML, implemented in PAML v4.9 (Yang 2007) using likelihood ratio tests. The genes with FDR < 0.1 were considered to be significant and selected as potential PSG candidates.

Acknowledgements This work was funded by National Natural Science Foundation of China (No. 31402035). We wish to thank Dr. Eric Lord for reading through the paper and giving valuable comments.

References

Adams MG, Johnson E (2010) Seasonal changes in the skin glands of Roe deer (*Capreolus capreolus*). Proc Zool Soc Lond 191:509–520

Adams M, Smith UM, Logan CV, Johnson CA (2008) Recent advances in the molecular pathology, cell biology and genetics of ciliopathies. J Med Genet 45:257–267

Bana NA, Nyiri A, Nagy J, Frank K, Nagy T, Steger V, Schiller M, Lakatos P, Sugar L, Horn P et al (2018) The red deer *Cervus elaphus* genome CerEla1.0: sequencing, annotating, genes, and chromosomes. Mol Genet Genomics 293:665–684

Chen L, Qiu Q, Jiang Y, Wang K, Lin Z, Li Z, Bibi F, Yang Y, Wang J, Nie W et al (2019) Large-scale ruminant genome sequencing provides insights into their evolution and distinct traits. Science. https://doi.org/10.1126/science.aav6202

Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 32:1792–1797

Fan Z, Li W, Jin J, Cui K, Yan C, Peng C, Jian Z, Bu P, Price M, Zhang X et al (2018) The draft genome sequence of forest musk deer (*Moschus berezovskii*). Gigascience. https://doi.org/10.1093/gigascience/giy038

Groves C (2007) Family cervidae. In: Prothero DR, Foss SE (eds) The evolution of artiodactyls. Johns Hopkins University Press, Baltimore

Grubb P (1993) Mammal species of the world. In: Wilson DE, Reeder DM (eds) A taxonomic geographic reference. Johns Hopkins University Press, Baltimore, pp 384–392

Han MV, Thomas GW, Lugo-Martinez J, Hahn MW (2013) Estimating gene gain and loss rates in the presence of error in genome assembly and annotation using CAFE 3. Mol Biol Evol 30:1987–1997

He Z, Zhang H, Gao S, Lercher MJ, Chen WH, Hu S (2016) Evolview v2: an online visualization and management tool for customized and annotated phylogenetic trees. Nucleic Acids Res 44:W236–W241

Hedges SB, Marin J, Suleski M, Paymer M, Kumar S (2015) Tree of life reveals clock-like speciation and diversification. Mol Biol Evol 32:835–845

Huang DW, Sherman BT, Lempicki RA (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc 4:44–57

King SM (2010) Sensing the mechanical state of the axoneme and integration of Ca^{2+} signaling by outer arm dynein. Cytoskeleton (Hoboken) 67:207–213

King SJ, Schroer TA (2000) Dynactin increases the processivity of the cytoplasmic dynein motor. Nat Cell Biol 2:20–24

Kumar S, Stecher G, Suleski M, Hedges SB (2017) TimeTree: a resource for timelines, timetrees, and divergence times. Mol Biol Evol 34:1812–1819

Lewin B, Cassimeris L, Plopper G (2007) Cells. Jones & Bartlett Learning, Sudbury

Li L, Stoeckert CJ Jr, Roos DS (2003) OrthoMCL: identification of ortholog groups for eukaryotic genomes. Genome Res 13:2178–2189

Li Z, Lin Z, Ba H, Chen L, Yang Y, Wang K, Qiu Q, Wang W, Li G (2017) Draft genome of the reindeer (*Rangifer tarandus*). Gigascience 6:1–5

Meyer W, Seegers U, Bock M (2008) Annual secretional activity of the skin glands in the Southern pudu (*Pudu puda* Molina 1782, Cervidae). Mamm Biol 73:392–395

Müller-Schwarze D, Altieri R, Porter N (1984) Alert odor from skin gland in deer. J Chem Ecol 10:1707–1729

Mykytowycz R, Goodrich BS (1974) Skin glands as organs of communication in mammals. J Invest Dermatol 62:124–131

Olson MV (1999) When less is more: gene loss as an engine of evolutionary change. Am J Hum Genet 64:18–23

Ralls K (1971) Mammalian scent marking. Science 171:443–449

Sahm A, Bens M, Platzer M, Szafranski K (2017) PosiGene: automated and easy-to-use pipeline for genome-wide detection of positively selected genes. Nucleic Acids Res 45:e100

Schrider DR, Hahn MW (2010) Gene copy-number polymorphism in nature. Proc Biol Sci 277:3213–3221

Stamatakis A (2014) RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30:1312–1313

Thiessen D, Rice M (1976) Mammalian scent gland marking and social behavior. Psychol Bull 83:505–539

Wood WF (2003) Volatile components in metatarsal glands of sika deer, *Cervus nippon*. J Chem Ecol 29:2729–2733

Yang Z (2007) PAML 4: phylogenetic analysis by maximum likelihood. Mol Biol Evol 24:1586–1591

Zhang Z, Li J, Zhao XQ, Wang J, Wong GK, Yu J (2006) KaKs_Calculator: calculating Ka and Ks through model selection and model averaging. Genomics Proteomics Bioinformatics 4:259–263

Zhang Z, Xiao J, Wu J, Zhang H, Liu G, Wang X, Dai L (2012) ParaAT: a parallel tool for constructing multiple protein-coding DNA alignments. Biochem Biophys Res Commun 419:779–781

Zhang C, Chen L, Zhou Y, Wang K, Chemnick LG, Ryder OA, Wang W, Zhang G, Qiu Q (2018) Draft genome of the milu (*Elaphurus davidianus*). Gigascience 7:1–6