Nucleotide Polymorphisms in the Canine *Noggin* Gene and Their Distribution Among Dog (*Canis lupus familiaris*) Breeds

**Yuji Ishii · Tatsuya Takizawa · Hiroshi Iwasaki ·
Yukihiro Fujita · Masaru Murakami ·
Jay C. Groppe · Kazuaki Tanaka**

Abstract   Noggin (NOG) is an important regulator for the signaling of bone morphogenetic proteins. In this study, we sequenced the complete coding sequence of the canine *NOG* gene and characterized the nucleotide polymorphisms. The sequence length varied from 717 to 729 bp, depending on the number of a 6-bp tandem repeat unit (GGCGCG), an insertion that has not been observed in other mammalian *NOG* genes investigated to date. It results in extensions of $(Gly-Ala)_{3-5}$ in the putative NOG protein. To survey the distribution of these tandem repeat polymorphisms, we analyzed 126 individuals in seven dog breeds. We identified only three alleles: $(GGCGCG)_3$, $(GGCGCG)_4$, and $(GGCGCG)_5$. Although the allele frequencies were remarkably different among the breeds, the three alleles were present in all seven of the breeds and did not show any deviation from Hardy–Weinberg equilibrium.

Y. Ishii · T. Takizawa · H. Iwasaki · K. Tanaka (✉)
Laboratory of Animal Biotechnology, School of Veterinary Medicine, Azabu University, Chuou-ku, Sagamihara 252-5201, Japan
e-mail: tanakak@azabu-u.ac.jp

Y. Fujita
Laboratory of Veterinary Surgery 2, School of Veterinary Medicine, Azabu University, Chuou-ku, Sagamihara 252-5201, Japan

M. Murakami
Laboratory of Molecular Biology, School of Veterinary Medicine, Azabu University, Chuou-ku, Sagamihara 252-5201, Japan

J. C. Groppe
Department of Biomedical Sciences, Baylor College of Dentistry, TAMHSC, Dallas, TX 75246, USA

Introduction

Noggin (NOG) is a member of the bone morphogenetic protein (BMP) family and a major extracellular antagonist of the BMP. NOG regulates dorsal induction and joint formation in embryonic development by blocking BMP signaling (Smith and Harland 1992; Zimmerman et al. 1996; Brunet et al. 1998; McMahon et al. 1998; Canalis et al. 2003). NOG also controls bone regeneration and homeostasis in adults (Canalis et al. 2003; Canalis 2009). The coding sequences of mammalian NOG genes are contained in a single exon and are highly conserved among species (Valenzuela et al. 1995). Nucleotide and amino acid sequence similarities are more than 90% among humans, mice, and other mammals (Valenzuela et al. 1995). In humans, several point mutations in the *NOG* gene have been linked to congenital skeletal malformations, for example, proximal symphalangism or multiple synostoses syndrome (Gong et al. 1999; Marcelino et al. 2001; Hirshoren et al. 2008). Therefore, the *NOG* gene is a potential candidate for involvement in congenital skeletal malformations in domesticated dogs. Partial DNA sequences at this locus have been defined by the Canine Genome Project (NW_876332; *Canis familiaris* chromosome 9 genomic contig). To the best of our knowledge, no nucleotide polymorphisms have been described in the canine *NOG* gene. In the present study, we first determined the DNA sequence of the complete coding sequence of the canine *NOG* gene, to scan for all the possible nucleotide polymorphisms encoded at this locus. We subsequently surveyed the allelic distribution of the variable number of tandem repeats (VNTRs) that were identified in the canine *NOG* gene.

Materials and Methods

Tissue and nail samples from non-sibling dogs were collected at veterinary hospitals and dog-grooming shops in Japan; 127 samples were obtained from 24 long-haired Miniature Dachshunds, 25 long-haired Chihuahuas, 23 Miniature Poodles, 19 Shih Tzus, 13 Papillons, 12 Malteses, 10 Yorkshire Terriers, and 1 Beagle. Genomic DNA was extracted from tissue or nails by a standard phenol–chloroform extraction method (Sambrook et al. 1989) or using the QuickGene DNA tissue kit (Fuji Film, Tokyo, Japan).

Amplification and Sequencing of Canine *NOG* Coding Sequence

Four animals from three breeds were selected for DNA sequence determination of the complete coding sequence for the canine *NOG* gene (one Beagle, two Miniature Dachshunds, and one Chihuahua). Primers for polymerase chain reaction (PCR) and sequencing were designed from dog genome draft sequences (chromosome 9, NW_876332; rat (NM_012990), mouse (NM_008711), and human (NM_005450) *NOG* gene sequences (Table 1). The primer sets of Can*NOG*-F1 and Can*NOG*-R1

Table 1 Primers for canine *NOG* gene

Primer	Sequence (5′ to 3′)	Position on NW_876332[a]
Can*NOG*-F1	TGGTGATGGAGCTGAAAGTG	13273923–13273942
Can*NOG*-R1	ACCACAGCCACATCTGTAACTTC	13274986–13275008
Can*NOG*-F2	ACCCGGACCCTATCTTTGAC	13274685–13274704
Can*NOG*-R2	TTTCTGGCTACAGAGACCTAGCT	13275418–13275440
Can*NOG*-S1[b]	GAAGTTACAGATGTGGCTGTGGT	Reverse complement of R1
Can*NOG*-S2[b]	GTCAAAGATAGGGTCCGGGT	Reverse complement of F2

[a] *Canis familiaris* chromosome 9 genomic contig, whole genome shotgun sequence

[b] Primers used only for sequencing

produce about 1.1 kbp DNA products, whereas the primer sets of Can*NOG*-F2 and Can*NOG*-R2 produce about 700 bp. The two products overlapped with one another in the canine *NOG* coding sequence.

The PCR was performed according to the manufacturer's instructions for PrimeStar HS DNA polymerase with GC buffer (Takara Biotechnology, Shiga, Japan). The program included first denaturation at 98°C for 2 min; 35 cycles of denaturation at 98°C for 20 s, annealing at 55°C for 5 s, and extension at 72°C for 60 s; and final extension at 72°C for 3 min. The amplified products of each reaction were cloned into the plasmid pCR-Blunt II-TOPO using the Zero Blunt TOPO PCR Cloning kit (Invitrogen, Carlsbad, USA). At least five clones were isolated from each individual and sequenced using the Big Dye Terminator version 3.1 Cycle Sequencing Kit and ABI 3100 sequencer (Applied Biosystems, Foster City, USA) with M13 universal primers and additional sequencing primers (Table 1).

Detection of the VNTR Polymorphisms in the *NOG* Coding Sequence

In order to survey the distribution of the VNTR alleles in dog breeds, we analyzed 126 animals from seven breeds (Table 2). The 300 bp DNA products of the canine *NOG* gene, including the 6-bp tandem repeats (GGCGCG), were PCR-amplified

Table 2 Genotype and allele frequency of $(GGCGCG)_n$ VNTR in canine *NOG* gene among seven dog breeds

Breed (number of animals sampled)	Genotype						Allele frequency		
	3	4	5	3/4	3/5	4/5	3	4	5
Miniature Dachshund, long hair (24)	1	7	0	7	7	2	0.333	0.479	0.188
Chihuahua, long hair (25)	3	0	6	5	7	4	0.360	0.180	0.460
Miniature poodle (23)	2	2	2	5	6	6	0.326	0.326	0.348
Shih Tzu (19)	9	0	2	3	4	1	0.658	0.105	0.237
Papillon (13)	5	0	0	2	4	2	0.615	0.154	0.231
Maltese (12)	2	1	0	0	6	3	0.417	0.208	0.375
Yorkshire Terrier (10)	8	0	0	1	1	0	0.900	0.050	0.050

using primer sets of Can*NOG*-F2 and Can*NOG*-R1 and then purified with the High Pure PCR Product Purification Kit (Roche Applied Science, Mannheim, Germany). Purified DNAs were sequenced directly with the primer Can*NOG*-F2 as described above. Heterozygous individuals that had a different number of tandem repeats were genotyped from their electropherograms.

Results and Discussion

In four domestic dogs, we sequenced the entire coding region of the *NOG* gene, which included an unidentified region of 91 nucleotides in NW_876332. These nucleotide sequences were entered into the DDBJ, EMBL, and GenBank nucleotide databases with accession numbers AB544074–AB544077. After multiple alignment of canine sequences with human *NOG* (NM_005450) and mouse *Nog* (NM_008711), we found that the coding sequence of the canine *NOG* gene is longer than that of humans and mice. The coding sequence of the human and mouse genes is consistently 699 bp in length. The coding sequence in the four dogs showed three polymorphisms, of 717, 723, and 729 bp, caused by inserted sequences of 18, 24, and 30 bp, respectively. These inserted sequences in the canine *NOG* gene were constructed by different numbers of the 6-bp tandem repeats (GGCGCG) unit (Fig. 1). These GGCGCG tandem repeats immediately preceded the inserted sequences. Thus, these sequence-length polymorphisms are the result of the duplicated insertion of this repeat unit. The insertions of tandem repeats were aligned with the reading frame of the *NOG* gene and found to encode a Gly–Ala repeat

Fig. 1 The coding region of the canine *NOG* gene and polymorphisms detected in this study. **a** Structure of the coding region of the canine *NOG* gene. The solid upper line indicates the partial sequence of canine chromosome 9 (NW_876332), and the broken line represents the undeciphered region in NW_876332. The start codon of canine *NOG* was located in the unidentified region, −91 bp from position 13274612 on NW_876332. The *white blocks* of the diagram represent the canine *NOG* coding sequence, and the *shaded block* indicates the region of the inserted sequence. **b** The nucleotide and amino acid sequences around the canine-specific insertion. The repeating units are ruled off by vertical lines. The numbers at the *bottom* indicate the number of repetitions. Repetition of GGCGCG was translated into the amino acids Gly–Ala

Fig. 2 Comparison of partial amino acid sequences of vertebrate Noggin. *Asterisks* indicate the conserved amino acids among the six species. The *gray box* encloses the polyglycine loop (Gly–Gly–Gly–Gly–Gly–Ala–Ala) insertion, which is present only in mammalian Noggin (Groppe et al. 2002). The *highlighted letters* indicate extension of the Gly–Ala repeat in the canine Noggin. The Gly–Ala repeat is connected to the polyglycine loop. Citation of sequences: human NM_005450, mouse NM_008711, pig NM_001143691, dog AF544077, chicken NM_204123, and *Xenopus* NM_00185644

(Fig. 1). These (Gly–Ala)$_{3–5}$ insertions are located in the polyglycine loop which is present only in mammalian Noggin (Fig. 2). The polyglycine loop is the most flexible part of the mammalian Noggin protein, as a complex with the BMP ligand (Groppe et al. 2002). Excluding these inserted-sequence regions, the coding sequence of the canine *NOG* gene displayed 96.7% sequence similarity with the nucleotides of human *NOG* and 94.0% with mouse *Nog*. No other polymorphic sites were found among the four canine *NOG* sequences, except the above-mentioned inserted sequences.

We sequenced part of the *NOG* gene, including the tandem repeat (GGCGCG)$_n$, in 126 dogs from seven breeds, and identified three alleles, (GGCGCG)$_3$, (GGCGCG)$_4$, and (GGCGCG)$_5$, among these individuals (Table 2). The genotypic and allelic frequencies of these VNTR alleles did not deviate from Hardy–Weinberg equilibrium among the seven breeds. Although the three alleles of the VNTR were present in all the dog breeds analyzed in this study, distributions varied widely among the breeds. (GGCGCG)$_4$ had the highest frequency in the Miniature Dachshund, whereas (GGCGCG)$_3$ was extremely high in the Shih Tzu, Papillon, and Yorkshire Terrier (Table 2). Since the three *NOG* VNTR alleles were distributed widely among various dog breeds, these sequence-length polymorphisms arose long before the modern dog breeds were established several hundred years ago.

Because we collected DNA samples from visibly healthy dogs, each VNTR allele in the canine *NOG* gene did not appear to be linked to fatal outcomes and severe pathological phenotypes in young dogs. Considering the important role of *NOG* in bone regeneration and homeostasis in adults (Canalis 2009), however, the three canine *NOG* variations theoretically could be associated with susceptibility to bone or joint disorders in the lifetime of the domestic dog. Nevertheless, the polymorphisms are most likely silent with respect to function given that (1) the additional residues encoded by the repeats were inserted in the most flexible portion of the Noggin protein, (2) the segment is well removed from the ligand-binding interface, (3) the loop is expanded in other mammals with no apparent loss of function, and (4) the inserted residues are small and functionally inert in the context of a disordered loop (Supplementary Fig. 1). In contrast, polymorphisms resulting in insertion of tracts of polyglutamine in transcription factors can have a toxic

gain-of-function effect in neurodegenerative disorders, triggering conformational changes that disrupt gene regulation (Riley and Orr, 2006) or culminate in formation of amyloid-like fibrils (Robertson et al. 2011). Indeed, insertion of glutamine repeats in Runx-2 protein, a transcription factor that regulates osteoblast differentiation, can have dramatic effects on canine skeletal morphogenesis (Fondon and Garner 2004). Introduction of amide-containing side-chains, with two functional groups capable of hydrogen bonding, can enhance interactions within multiprotein transcriptional complexes, leading to upregulation of target genes involved in skeletogenesis. Therefore, although polymorphisms within domestic dog breeds can have pronounced effects on development, the structural context and nature of the side-chain of the inserted residues appears to dictate whether the effect is functionally silent, accumulating over evolutionarily large time spans, or pronounced, capable of generating rapid diversity of form.

References

Brunet LJ, McMahon JA, McMahon AP, Harland RM (1998) Noggin, cartilage morphogenesis, and joint formation in the mammalian skeleton. Science 280:1455–1457

Canalis E (2009) Growth factor control of bone mass. J Cell Biochem 108:769–777

Canalis E, Economides AN, Gazzerro E (2003) Bone morphogenetic proteins, their antagonists, and the skeleton. Endocr Rev 24:218–235

Fondon JW, Garner HR (2004) Molecular origins of rapid and continuous morphological evolution. Proc Natl Acad Sci USA 101:18058–18063

Gong Y, Krakow D, Marcelino J, Wilkin D, Chitayat D, Babul-Hirji R, Hudgins L, Cremers CW, Cremers FP, Brunner HG, Reinker K, Rimoin DL, Cohn DH, Goodman FR, Reardon W, Patton M, Francomano CA, Warman ML (1999) Heterozygous mutations in the gene encoding noggin affect human joint morphogenesis. Nature Genet 21:302–304

Groppe J, Greenwald J, Wiater E, Rodriguez-Leon J, Economides AN, Kwiatkowski W, Affolter M, Vale WW, Belmonte JC, Choe S (2002) Structural basis of BMP signalling inhibition by the cystine knot protein Noggin. Nature 420:636–642

Hirshoren N, Gross M, Banin E, Sosna J, Bargal R, Raas-Rothschild A (2008) P35S mutation in the *NOG* gene associated with Teunissen-Cremers syndrome and features of multiple *NOG* joint-fusion syndromes. Eur J Med Genet 51:351–357

Marcelino J, Sciortino CM, Romero MF, Ulatowski LM, Ballock RT, Economides AN, Eimon PM, Harland RM, Warman ML (2001) Human disease-causing *NOG* missense mutations: effects on noggin secretion, dimer formation, and bone morphogenetic protein binding. PNAS 98:11353–11358

McMahon JA, Takada S, Zimmerman LB, Fan CM, Harland RM, McMahon AP (1998) Noggin-mediated antagonism of BMP signaling is required for growth and patterning of the neural tube and somite. Genes Dev 12:1438–1452

Riley BE, Orr HT (2006) Polyglutamine neurodegenerative diseases and regulation of transcription: assembling the puzzle. Genes Dev 20:2183–2192

Robertson AL, Bate MA, Androulakis SG, Bottomly SP, Buckle AM (2011) PolyQ: a database describing the sequence and domain context of polyglutamine repeats in proteins. Nucleic Acids Res 39(Database issue):D272–D276

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual, 2nd edn. Cold Spring Harbor Laboratory Press, New York

Smith WC, Harland RM (1992) Expression cloning of noggin, a new dorsalizing factor localized to the Spemann organizer in *Xenopus* embryos. Cell 70:829–840

Valenzuela DM, Economides AN, Rojas E, Lamb TM, Nunez L, Jones P, Lp NY, Espinosa R 3rd, Brannan CI, Gilbert DJ, Copeland NG, Jenkins NA, Le Beau MM, Harland RM, Yancopoulos GD (1995) Identification of mammalian noggin and its expression in the adult nervous system. J Neurosci 15:6077–6084

Zimmerman LB, De Jesus-Escobar JM, Harland RM (1996) The Spemann organizer signal noggin binds and inactivates bone morphogenetic protein 4. Cell 86:599–606