Array-based comparative genomic hybridization–guided identification of reference genes for normalization of real-time quantitative polymerase chain reaction assay data for lymphomas, histiocytic sarcomas, and osteosarcomas of dogs

Pei-Chien Tsai, PhD, and Matthew Breen, PhD

Objective—To identify suitable reference genes for normalization of real-time quantitative PCR (RT-qPCR) assay data for common tumors of dogs.

Sample—Malignant lymph node (n = 8), appendicular osteosarcoma (9), and histiocytic sarcoma (12) samples and control samples of various nonneoplastic canine tissues.

Procedures—Array-based comparative genomic hybridization (aCGH) data were used to guide selection of 9 candidate reference genes. Expression stability of candidate reference genes and 4 commonly used reference genes was determined for tumor samples with RT-qPCR assays and 3 software programs.

Results—*LOC611555* was the candidate reference gene with the highest expression stability among the 3 tumor types. Of the commonly used reference genes, expression stability of *HPRT* was high in histiocytic sarcoma samples, and expression stability of *Ubi* and *RPL32* was high in osteosarcoma samples. Some of the candidate reference genes had higher expression stability than did the commonly used reference genes.

Conclusions and Clinical Relevance—Data for constitutively expressed genes with high expression stability are required for normalization of RT-qPCR assay results. Without such data, accurate quantification of gene expression in tumor tissue samples is difficult. Results of the present study indicated *LOC611555* may be a useful RT-qPCR assay reference gene for multiple tissue types. Some commonly used reference genes may be suitable for normalization of gene expression data for tumors of dogs, such as lymphomas, osteosarcomas, or histiocytic sarcomas. (*Am J Vet Res* 2012;73:1335–1343)

Cancer is the most common cause of death of dogs. Twenty-three percent of dogs overall[1] and 45% of dogs ≥ 10 years old[2] die of cancer. Reported annual incidence rates for cancer in dogs are between 310 and 958/100,000 dogs.[3–5] Although the biological behavior of many types of tumors in dogs is similar to that of tumors in humans, the molecular pathways that cause development of tumors in dogs are not well understood. Other authors have identified a complete high-quality canine genome sequence,[6] which may be useful for investigation of the molecular biology of tumors of dogs. Determination of gene expression profiles may aid understanding of cell regulatory processes in tumor cells and assist in identification of genes that regulate tumorigenesis and biological behavior of tumors. Quantitative information regarding gene expression can be obtained via RT-qPCR assay, which enables simultaneous measurement of expression of select genes in multiple assay samples.[7–9] Although RT-qPCR is the most accurate, sensitive, and reproducible method for determination of gene expression, results for genes of interest must

Abbreviations

aCGH	Array-based comparative genomic hybridization
C_T	Cycle threshold
NF	Normalization factor
RT-qPCR	Real-time quantitative PCR

Received January 7, 2011.
Accepted September 12, 2011.
From the Department of Molecular Biomedical Science, College of Veterinary Medicine (Tsai, Breen), and the Center for Comparative Medicine and Translational Research (Breen), North Carolina State University, Raleigh, NC 27606; and the Cancer Genetics Program, UNC Lineberger Comprehensive Cancer Center, University of North Carolina, Chapel Hill, NC 27599 (Breen).
Supported in part by grants from the AKC Canine Health Foundation and the National Institutes of Health (awarded to Dr. Breen).
The authors thank Drs. Rachael Thomas, Andrea Angstadt, and Benoit Hedan for providing data for DNA copy number differences among canine lymphoma, osteosarcoma, and histiocytic sarcoma samples, respectively, and Drs. Steven Suter and Kristy Richards for providing some of the canine tissues and RNA samples.
Address correspondence to Dr. Breen (Matthew_Breen@ncsu.edu).

be normalized relative to results for other genes. Gene expression results are typically normalized relative to expression of housekeeping or constitutively expressed reference genes, which should have consistent expression in tissues regardless of environmental conditions.[10] However, the commonly used reference genes glyceraldehyde-3-phosphate dehydrogenase (*GAPDH*), beta actin (*ACTB*), and beta-2-microglubulin (*B2M*) may not be consistently expressed in all tissue types and for all diseases.[11,12] Therefore, normalization of gene expression results relative to expression of these reference genes can cause inaccurate assay results and study conclusions. Consequently, use of reference genes with consistent expression may improve accuracy and sensitivity of RT-qPCR assays.

Array-based comparative genomic hybridization analysis is a method for identification of DNA segment copy number aberrations (ie, gene dosage) in genomic DNA isolated from malignant cells.[13–22] Several consistent DNA copy number aberrations have been identified in samples of lymphomas, osteosarcomas, and histiocytic sarcomas of dogs. Regions of the canine genome that have consistent DNA copy numbers among these types of tumors have also been identified. Expression stability may be high for genes in these genomic regions. Such genes may be useful as reference genes for normalization of gene expression RT-qPCR results for tumors.

Objectives of the study reported here were to compare the aCGH profiles of 3 types of tumors (lymphoma, osteosarcoma, and histiocytic sarcoma) obtained from dogs, identify candidate reference genes with consistent copy number and expression stability among those tumor types, and compare results for candidate reference genes with those for genes (hypoxanthine-phosphoribosyl transferase [*HPRT*], ATP-synthase subunit 5B [*ATP5B*], ribosomal protein L32 [*RPL32*], and ubiquitin [*Ubi*]) determined by other authors[23–25] to have stable expression in canine tissues.

Materials and Methods

Sample—Tumor samples (lymphoma [n = 8], appendicular osteosarcoma [9], and histiocytic sarcoma [12]) were obtained from client-owned dogs admitted to the Veterinary Medical Center at the College of Veterinary Medicine at North Carolina State University. Tumor samples were acquired from dogs prior to initiation of chemotherapy or radiotherapy, in accordance with protocols that were reviewed and approved by the appropriate Institutional Review Board and Institutional Animal Care and Use Committee. Informed consent was obtained from clients for use of tumor samples. Samples of lymph node (n = 2), spleen (1), lung (1), and bone (1) unaffected by tumors that had been obtained previously from 3 mixed-breed dogs were used for determination of RT-qPCR assay efficiency and served as control samples for each type of tumor. Immediately following surgical excision of tumors, half of a representative sample of each tumor was fixed in neutral-buffered 10% formalin. The other half of each sample of tumors was snap-frozen and kept in liquid nitrogen until RNA extraction for RT-qPCR assay and determination of gene expression stability via 3 software[a–c]-based methods.[26–28]

Histologic evaluation of tumor samples—Tumor samples obtained from dogs were submitted for evaluation by personnel of the North Carolina State University College of Veterinary Medicine pathology service. Tumor samples were examined after H&E and immunohistocytochemical (as needed) staining to determine tumor types.

Selection of candidate reference genes—A 1-Mb resolution genome assembly-integrated microarray[18] was used by other investigators to determine aCGH (ie, cytogenetic profile) data for various tumor samples (lymphoma [n = 122],[29] appendicular osteosarcoma [123],[30] and histiocytic sarcoma [86][31]). Those aCGH data were determined by calculation of the frequency of copy number neutrality (ie, 2 copies of loci detected) for each of 2,097 genomic loci in the microarray. Genomic regions with consistent copy number stability (ie, > 90% of tumor samples tested did not have DNA copy number changes) were identified for evaluation in the present study. Candidate reference genes for normalization of RT-qPCR gene expression data were selected from genomic regions determined by those other authors[29–31] to have stable expression in tumor samples. Gene sequences were evaluated to ensure they were not known to be involved in development of lymphoma, osteosarcoma, or histiocytic sarcoma. Four genomic regions were identified in which 9 candidate reference genes were selected. Genomic locations and cytogenetic

Table 1—Genomic regions with high copy number neutrality among canine lymphoma (n = 122), histiocytic sarcoma (86), and osteosarcoma (123) samples* and genes identified in those regions.

Clone address‡	Canine chromosome	Genomic location§	Copy number neutrality (%)†			Genes in genomic region
			Lymphoma	Histiocytic sarcoma	Osteosarcoma	
326-J08	7	76,288,559–76,478,148	98.1	94.1	90.7	*LAMA1*
313-F22	9	36,694,888–36,863,023	98.1	91.0	92.6	*PPM1E* and *TRIM37*
315-P19	25	6,818,798–6,993,295	98.1	93.2	90.6	*SMAD9*
122-H22	26	19,737,413–19,909,922	98.8	90.8	96.7	*LE10*, *LOC611555*, *SPPL3*, *HNF-1*, and *OASL*

*Results were determined via aCGH analysis of data reported for lymphoma,[29] histiocytic sarcoma,[31] and osteosarcoma[30] samples obtained from dogs. †Numbers are percentage of loci with a neutral number (2) of copies among tumor samples. Copy numbers for each of 2,097 genomic loci in tumor samples were determined with a microarray.[18] Genomic locations with consistent copy number neutrality (ie, > 90% of tumor samples tested did not have DNA copy number changes) were selected for further evaluation of gene expression stability. ‡Clone address is the address of the genomic region in a canine genome library.[e] §Numbers are the DNA base locations of the sequence in a *Canis lupus familiaris* genome assembly.[d]

Table 2—Characteristics and RT-qPCR assay metrics of genes identified as candidates for use as references for normalization of RT-qPCR assay C_T data for canine lymphoma, histiocytic sarcoma, and osteosarcoma samples.

Gene description	Gene symbol	Accession No.*	Function of gene product	Primer sequences (5′–3′)	RT-qPCR product length (bp)	R^2	RT-qPCR efficiency
Laminin α1	*LAMA1*	XM_537324.2	Component of glycoproteins; cell differentiation and migration	F: GCACAACACCACGGGGGACC R: AGGTGGCAGGTGGGGCTGAA	141	0.993	100.4
Protein phosphatase, Mg^{2+}/Mn^{2+} dependent, 1E	*PPM1E*	XM_848160	Protein phosphatase; CaM inactivation	F: AAACAGATGGCACAGAAGGG R: TTTTTGATGGCATGGATTGA	160	0.998	95.8
Tripartite motif-containing 37	*TRIM37*	XM_537697.2	Peroxisomal protein	F: CAGAGCTCCCTGACTTGGAC R: AATGCTCTCCACGCTCTGTT	156	1.000	95.8
SMAD family member 9	*SMAD9*	XM_852737.1	SMAD family polypeptide; cell signaling	F: GAGAGCCCCTATCAACACTCAGACT R: CGGGAGGATGCCTGGAACGTC	122	0.990	100.9
60S acidic ribosomal protein P0	*L10E/RPLP0*	XM_846329.1	60S ribosomal protein	F: CTTCCCACTTGCTGAAAAGG R: TGTCCGATTCCAACTCTTCC	149	0.990	99.4
Hypothetical protein LOC611555	*LOC611555*	XM_849238.1	Uncharacterized protein	F: GCCTGGGGCTTGGAGCAGTG R: TGGGCTCGGAATTCGGGGGT	158	0.999	96.8
Signal peptide peptidase 3	*SPPL3*	XM_543427.2	Intramembrane cleaving protease	F: CGACCGTGGCATCCCGCATT R: GGCTCAGACCACATCCGCCG	127	0.999	100.5
HNF1 homeobox A	*HNF-1*	XM_543429	Transcriptional activator for liver-specific genes	F: GCCCAGAGCCCCTTCATGGC R: AAGACCTGCTTGGTGGGCGT	176	0.993	100.4
2′-5′-oligoadenylate synthetase-like	*OASL*	NM_001048093	Viral RNA degradation	F: ACACCGCAGATCAATCATCA R: ACACCGCAGATCAATCATCA	188	0.997	99.2
ATP synthase, H^+ transporting, mitochondrial F1 complex β polypeptide	*ATP5B*	NM_001686	Mitochondrial ATP synthase subunit; ATP synthesis	F: GCACGGAAAATACAGCGTTT R: TTGCCACAGCTTCTTCAATG	187	0.995	100.8
Hypoxanthine phosphoribosyltransferase 1	*HPRT*	NM_000194	Enzyme; purine metabolism	F: TGCTCGAGATGTGATGAAGG R: TCCCCTGTTGACTGGTCATT	192	0.995	100.2
Polyubiquitin	*Ubi*	NM_001009202	Labeling proteins for proteasomal protein degradation	F: TCTTCGTGAAAACCCTGACC R: CCTTCACATTCTCGATGGTG	305	0.998	97.3
60S ribosomal protein L32	*RPL32*	XM_540107	Ribosomal protein	F: ATGCCCAACATTGGTTATGG R: CTCTTTCCACGATGGCTTTG	180	1.000	100.6

For determination of RT-qPCR assay metrics (R^2 and RT-qPCR efficiency), an RT-qPCR assay was performed for five 10-fold dilutions of a pool of RNA samples isolated from nonneoplastic canine lymph node samples.[8]

*Accession number of gene in a database[j] of nucleotide sequences.

CaM = Ca^{2+}/calmodulin-dependent protein kinase. F = Forward. R = Reverse.

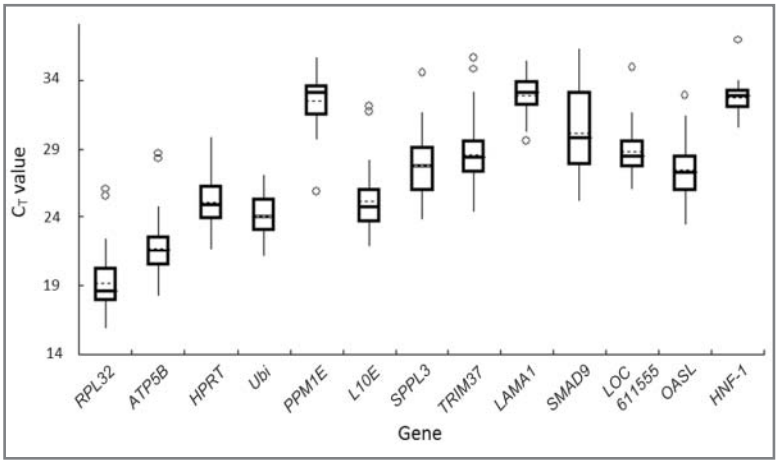

Figure 1—Box-and-whiskers plot of RT-qPCR assay C_T values for various genes in canine tumor (lymphoma [n = 8], osteosarcoma [9], and histiocytic sarcoma [12]) samples. In each box, the central solid horizontal line represents the median value, the central dotted horizontal line represents the mean value, and bottom and top of the boxes represent the 25th and 75th percentiles, respectively. Whiskers represent the range of values excluding outliers, and white circles represent outlier values. An outlier is defined as a value that is less than the 25th percentile value minus 1.5 times the interquartile range or a value that is greater than the 75th percentile value plus 1.5 times the interquartile range.

data for these genomic regions and candidate reference genes were determined with a database[d] and summarized (**Table 1**[e]).

RNA extraction—Tumor sample RNA was isolated by use of an RNA isolation kit[f] with genomic DNA digestion[g] in accordance with manufacturer's instructions. Total RNA concentration was determined with a spectrophotometer,[h] and quality and integrity of RNA were determined with a microfluidics-based analyzer.[i] All RNA samples used in the study had an RNA integrity value > 7.0.

RT-qPCR assay—Primer sequences were determined for each candidate reference gene by use of available transcript sequences[j] and primer design software.[32,k] Reverse transcription was performed with a reverse transcription kit[l] in accordance with manufacturer's instructions. Briefly, 1 µg of isolated RNA was incubated with genomic DNA degradation buffer at 42°C for 2 minutes. Reverse transcription primer

mix, reverse transcription buffer, and reverse transcriptase were added to each RNA sample, and samples were incubated at 42°C for 15 minutes. Reverse transcriptase was inactivated by incubation at 95°C for 3 minutes. The resultant cDNA samples were stored at –20°C until analysis.

The RT-qPCR assays were performed for triplicate cDNA samples with a 96-well format PCR kit[m] and thermal cycler.[n] The RT-qPCR assay mixtures comprised 1X asymmetric cyanine dye PCR mixture, 1µM forward and reverse primers, 25 ng of template cDNA, and 0.5 µL of fluorescein passive reference dye[o]; sterile water was added to achieve a final volume of 25 µL. The RT-qPCR assay amplification conditions comprised 1 cycle at 95°C for 5 minutes and 40 cycles at 95°C for 10 seconds and 60°C for 30 seconds. A melting curve was generated by heating the samples from 55° to 95°C in 1°C increments with a 15-second hold at each temperature. A control sample containing no cDNA template and another control sample containing no reverse transcriptase were used as controls for sample contamination and genomic contamination, respectively. None of the control sample reactions had detectable products. Amplified RT-qPCR assay products were separated by electrophoresis on 2% agarose gels, and nucleic acids were stained.[p] For determination of the RT-qPCR assay efficiency for each gene evaluated, an RT-qPCR assay was performed for five 10-fold dilutions of a pool of RNA samples isolated from nonneoplastic canine lymph node tissue.[8] Genes and corresponding primer sets used in the RT-qPCR assay were summarized (**Table 2**).

Reference gene expression stability—Expression of candidate reference genes in tumor samples was quantified via determination of the number of RT-qPCR assay cycles required to attain a threshold value in the exponential phase of the PCR reaction. Data were exported to a spreadsheet,[q] and 3 computer programs[a–c] were used to determine expression stability of candidate reference genes on the basis of mathematical algorithms.

One of the programs (program 1[a]) was used to determine gene expression stability on the basis of the principle that the expression of 2 theoretically ideal reference genes will always be identical among tissue samples, regardless of experimental conditions.[26] Therefore, the program defined stability as the mean pairwise variation in expression of a particular reference gene, compared with the mean pairwise variation in expression of each of the other reference genes. Genes with low expression stability values had high gene expression stability among tumor samples (ie, the gene expression stability value was inversely related to gene expression stability). Via stepwise removal of the gene with the least expression stability and recalculation of the stability value, the genes with the highest expression stability (lowest stability values) were identified. The optimal number of genes for normalization of RT-qPCR data was also determined for each tumor type with program 1. For tissue types for which no single optimal reference gene has been identified, it is recommended that data be normalized against > 1 reference gene and that an NF be calculated.[11,26–28] Variation in the mean expression stability values among multiple reference genes is typically smaller than the variation for 1 reference gene. However, the number of reference genes used is a compromise between practical considerations (ie, convenience of performing experiments) and accuracy of the NF.[26,27] The NFs were calculated on the basis of the geometric means of expression of the reference genes with high expression stability. The pairwise variation between 2 sequential NFs (NF_n and NF_{n+1}, where NF_n is the geometric mean of the expression for the n-ranked reference gene and NF_{n+1} is the geometric mean of the expression for the n + 1–ranked reference gene) was calculated[23–26] to determine the benefit of adding extra reference genes for the normalization process.

Another program (program 2[b]) was used to determine gene expression stability via a model-based approach that estimated variation in expression among all candidate reference genes and subgroups of reference genes.[27] This feature made results robust against the effects of results for coexpressed candidate reference genes. For program 2 results, low gene expression stability values indicated high gene expression stability.

Another program (program 3[c]) was used to determine gene expression stability via ranking of candidate reference genes in accordance with variation in C_T values among tumor samples.[28] Therefore, SDs of C_T values were used as a measure of gene expression stability. Low SD of C_T values indicated high gene expression stability. Genes with an SD of C_T value > 1 were considered to have inconsistent expression (ie, low expression stability).

Results

RT-qPCR metrics—The RT-qPCR assay metrics for candidate and commonly used reference genes were summarized (Table 2). To ensure results of RT-qPCR assays could be compared, PCR efficiency for each gene was determined via a dilution method.[8] Each RT-qPCR assay had an efficiency value between 95% and 101%. All RT-qPCR assays had a single distinctive melt-curve analysis peak and a single amplicon of the expected size as determined via agarose gel electrophoresis.

Table 3—Genes identified as candidates for use as references for normalization of RT-qPCR assay C_T data listed in descending order of expression stability in canine lymphoma (n = 8), histiocytic sarcoma (12), and osteosarcoma (9) samples as determined by use of a program[a] algorithm on the basis of pairwise variation in expression of a gene, compared with that for each other gene.

Lymphoma	Histiocytic sarcoma	Osteosarcoma
LOC611555 and *HNF-1**	*LOC611555* and *OASL**	*HPRT* and *Ubi**
LAMA1	*SPPL3*	*L10E*
OASL	*HPRT*	*RPL32*
Ubi	*TRIM37*	*LOC611555*
SPPL3	*L10E*	*OASL*
HPRT	*RPL32*	*TRIM37*
ATP5B	*ATP5B*	*SPPL3*
RPL32	*SMAD9*	*ATP5B*
L10E	*Ubi*	*SMAD9*
SMAD9	*HNF-1*	*LAMA1*
PPM1E	*PPM1E*	*HNF-1*
TRIM37	*LAMA1*	*PPM1E*

*Expression stability of the 2 highest-ranked genes cannot be differentiated because the method calculates expression stability values on the basis of ratios of results for 2 genes.

Table 4—Genes identified as candidates for use as references for normalization of RT-qPCR assay C_T data listed in descending order of expression stability in canine lymphoma (n = 8), histiocytic sarcoma (12), and osteosarcoma (9) samples as determined by use of a program[b] ranking expression stability on the basis of results of an algorithm for estimation of variation in expression of each candidate reference gene among all samples and among various subgroups of samples.

Lymphoma		Histiocytic sarcoma		Osteosarcoma	
Gene	Expression stability	Gene	Expression stability	Gene	Expression stability
LOC611555	0.231	*HPRT*	0.172	*LOC611555*	0.185
RPL32	0.244	*LOC611555*	0.244	*RPL32*	0.324
HNF-1	0.346	*Ubi*	0.281	*Ubi*	0.344
LAMA1	0.357	*RPL32*	0.326	*HPRT*	0.361
OASL	0.408	*SPPL3*	0.412	*OASL*	0.370
HPRT	0.426	*TRIM37*	0.446	*SPPL3*	0.384
TRIM37	0.495	*ATP5B*	0.499	*TRIM37*	0.391
SPPL3	0.563	*L10E*	0.506	*L10E*	0.462
L10E	0.573	*OASL*	0.530	*ATP5B*	0.507
PPM1E	0.614	*PPM1E*	0.575	*SMAD9*	0.508
ATP5B	0.630	*HNF-1*	0.675	*LAMA1*	0.577
Ubi	0.658	*SMAD9*	0.848	*HNF-1*	0.617
SMAD9	0.810	*LAMA1*	1.001	*PPM1E*	0.904

Data are gene expression stability values. Low expression stability values indicate high gene expression stability.

Expression of candidate reference genes—The RT-qPCR assay C_T values for 13 genes (4 commonly used reference genes and 9 candidate reference genes identified in the present study) were summarized (**Figure 1**). Genes were allocated to 1 of 2 groups on the basis of median C_T values. Nine genes had high expression (median C_T value < 30), including all 4 of the commonly used reference genes (*RPL32*, *ATP5B*, *HPRT*, and *Ubi*) and 5 of the candidate reference genes identified in the present study (*L10E*, *SPPL3*, *TRIM37*, *LOC611555*, and *OASL*). Four genes (*PPM1E*, *LAMA1*, *SMAD9*, and *HNF-1*) had low expression (median C_T value > 30), all of which were candidate reference genes identified in the present study. The genes with the lowest range of C_T values among tumors (defined as the difference between the highest and lowest C_T values excluding outliers for a gene among all tumor samples) were *HNF-1* (C_T value range, 1.3), *LAMA1* (C_T value range, 1.6), and *LOC611555* (C_T value range, 1.9).

Gene expression stability among types of tumors—Genes were ranked on the basis of expression stability results as determined with programs 1, 2, and 3 (**Tables 3–5**). Expression stability rank order of each gene as determined with the 3 software programs and mean rank values were summarized (**Table 6**).

For lymphoma samples, *LOC611555* was determined to have 1 of the 2 highest expression stabilities by use of programs 1 and 2 (Tables 3 and 4) and the second highest expression stability by use of program 3 (Table 5). Results determined by use of all 3 programs indicated *PPM1E*, *SMAD9*, and *TRIM37* had the lowest expression stabilities. Ranking of genes on the basis of expression stability differed among the 3 programs, although agreement seemed to be good for those genes with the highest and lowest expression stabilities.

For histiocytic sarcoma samples, results obtained with the 3 programs differed regarding ranking of the genes with the highest expression stabilities (Tables 3–5). Expression stability of *LOC611555* was ranked as 1 of the 2 highest by use of program 1, second highest by use of program 2, and third highest by use of program 3. Expression stability of *HPRT* was ranked highest with program 2 but was ranked fourth highest with programs 1 and 3. Expression stability of *OASL* was ranked highest with program 3 and was ranked as 1 of the 2 highest with program 1. However, expression stability of *OASL* was ranked ninth highest with program 2. On the basis of the mean rank order of gene expression stability results of the 3 programs (Table 6), the genes with the highest expression stabilities in histiocytic sarcoma samples were *LOC611555*, *HPRT*, and *OASL*, whereas *PPM1E*, *LAMA1*, and *SMAD9* had the lowest expression stabilities.

For osteosarcoma samples, *LOC611555* was identified as the gene with the highest expression stability by use of programs 2 and 3 (Tables 4 and 5) but was identified as the gene with the fifth highest expression stability by use of program 1 (Table 3). Expression stability of *Ubi* was ranked as 1 of the 2 highest with

Table 5—Genes identified as candidates for use as references for normalization of RT-qPCR assay C_T data listed in descending order of expression stability in canine lymphoma (n = 8), histiocytic sarcoma (12), and osteosarcoma (9) samples as determined with a spreadsheet software[a]–based program[c] on the basis of SDs of C_T values calculated by use of pairwise correlations.

Lymphoma		Histiocytic sarcoma		Osteosarcoma	
Gene	SD_{CT}	Gene	SD_{CT}	Gene	SD_{CT}
HNF-1	0.26	*OASL*	0.53	*LOC611555*	0.59
LOC611555	0.43	*HNF-1*	0.54	*RPL32*	0.65
LAMA1	0.44	*LOC611555*	0.73	*PPM1E*	0.66
OASL	0.76	*HPRT*	0.85	*Ubi*	0.72
Ubi	0.87	*TRIM37*	0.91	*HNF-1*	0.80
RPL32	0.98	*PPM1E*	0.96	*HPRT*	0.81
L10E	1.00	*ATP5B*	0.98	*ATP5B*	0.82
SPPL3	1.10	*Ubi*	0.99	*LAMA1*	0.82
ATP5B	1.13	*SPPL3*	1.12	*L10E*	0.85
HPRT	1.19	*RPL32*	1.14	*TRIM37*	0.86
PPM1E	2.04	*L10E*	1.17	*SPPL3*	0.95
TRIM37	2.15	*LAMA1*	1.25	*OASL*	0.99
SMAD9	2.17	*SMAD9*	1.71	*SMAD9*	1.07

Low SD of C_T values indicate high gene expression stability.
SD_{CT} = Standard deviation of C_T values.

Table 6—Mean rank order of candidate genes for use as references for normalization of RT-qPCR assay C_T data for canine lymphoma (n = 8), histiocytic sarcoma (12), and osteosarcoma (9) samples on the basis of expression stability values determined by use of the 3 programs in Tables 3, 4, and 5.

Variable	*RPL32*	*ATP5B*	*HPRT*	*Ubi*	*PPM1E*	*L10E*	*SPPL3*	*TRIM37*	*LAMA1*	*SMAD9*	*LOC611555*	*OASL*	*HNF-1*
Lymphoma													
Program 1	9	8	7	5	12	10	6	13	3	11	1.5	4	1.5
Program 2	2	11	6	12	10	9	8	7	4	13	1	5	3
Program 3	6	9	10	5	11	7	8	12	3	13	2	4	1
Mean rank	5.7	9.3	7.7	7.3	11	8.7	7.3	10.7	3.3	12.3	1.5	4.3	1.8
Histiocytic sarcoma													
Program 1	7	8	4	10	12	6	3	5	13	9	1.5	1.5	11
Program 2	4	7	1	3	10	8	5	6	13	12	2	9	11
Program 3	10	7	4	8	6	11	9	5	12	13	3	1	2
Mean rank	7	7.3	3	7	9.3	8.3	5.7	5.3	12.7	11.3	2.2	3.8	8
Osteosarcoma													
Program 1	4	9	1.5	1.5	13	3	8	7	11	10	5	6	12
Program 2	2	9	4	3	13	8	6	7	11	10	1	5	12
Program 3	2	7	6	4	3	9	11	10	8	13	1	12	5
Mean rank	2.7	8.3	3.8	2.8	9.7	6.7	8.3	8	10	11	2.3	7.7	9.7

Values are rank order of expression stability of genes on the basis of results determined by use of program 1,[a] program 2,[b] and program 3.[c] Mean rank is the mean of the ranks assigned by use of programs 1, 2, and 3 for expression stability of genes in samples of each tumor type.

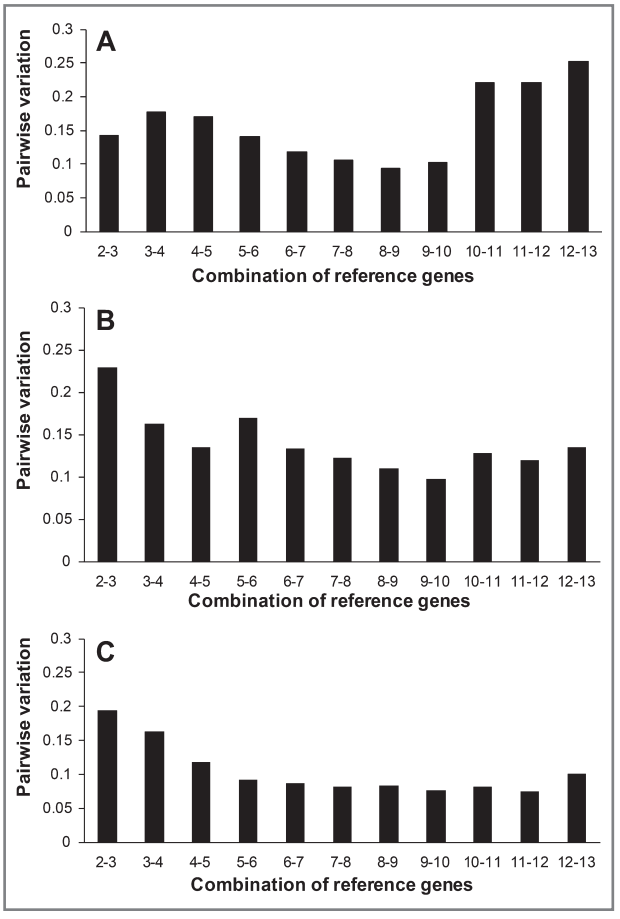

Figure 2—Results of NF pairwise variation analysis for determination of the optimal number of reference genes for normalization of RT-qPCR assay data for canine lymphoma (n = 8; A), histiocytic sarcoma (12; B), and osteosarcoma (9; C) samples obtained from dogs. Results were determined by use of a software program[a] to calculate[26] pairwise variation between the NFs for genes with sequential expression stability ranks. For gene combinations with a NF pairwise variation value < 0.15, inclusion of additional reference genes was considered unnecessary (ie, a pairwise variation cutoff of 0.15 was used to determine the optimal number of reference genes).

program 1, the third highest with program 2, and the fourth highest with program 3. Expression stability of *RPL32* was ranked second highest with programs 2 and 3 and fourth highest with program 1. On the basis of the mean rank order of gene expression stability results of the 3 programs (Table 6), *LOC611555*, *RPL32*, and *Ubi* had the highest expression stabilities and *SMAD9*, *LAMA1*, *PPM1E*, and *HNF-1* had the lowest expression stabilities in osteosarcoma samples.

Comparison of results among tumor types—Among the 3 tumor types, *LOC611555* was one of the most stably expressed and *SMAD9* and *PPM1E* were among the least stably expressed genes (Table 6). Among commonly used reference genes, *HPRT* had high expression stability in histiocytic sarcoma samples and *Ubi* and *RPL32* had high expression stability in osteosarcoma samples, but none of these genes were consistently identified as having the highest expression stability among all 3 tumor types.

Number of reference genes required for RT-qPCR data normalization—To determine the number of reference genes that should be used for normalization of RT-qPCR assay data for each tumor type, the contribution of expression stability results for each gene to the variation in NF was determined with program 1; results indicated the effects of adding additional reference genes. An NF pairwise variation cutoff of 0.15 was used; additional reference genes were not required if pairwise variation was below this value for a pair of candidate reference genes. For example, for lymphoma samples, the pairwise variation between the NFs of the 2 first-ranked and the 3 first-ranked reference genes was < 0.15 (**Figure 2**)[26,a]; therefore, there was no need to include a third gene for normalization of RT-qPCR data for this tumor type (ie, the optimal number of reference genes for normalization of lymphoma RT-qPCR data was determined to be 2). By use of expression stability rankings determined by use of program 1 (Table 3), the 2 reference genes with the highest expression stability (*LOC611555* and *HNF-1*) were identified for

use in normalization of RT-qPCR data for lymphoma samples. However, results indicated 4 genes were required for normalization of data for histiocytic sarcoma and osteosarcoma samples.

Discussion

Methods for selection of reference genes for normalization of gene expression data have been evaluated in other studies, but most of those studies involved evaluation of tissue samples obtained from humans. There is a need for identification of RT-qPCR assay reference genes suitable for use in studies in which medical conditions of dogs are investigated, particularly for those studies in which multiple tissue types or treatments are evaluated. Several methods have been used to determine suitability of commonly used and newly identified genes for normalization of RT-qPCR assay data, including determination of gene transcription profiles via microarray analysis.[23–25,33] However, use of such methods has not enabled identification of reference genes with consistent expression stability among tissue types and diseases. In the present study, we used 1-Mb resolution aCGH data determined by other authors[29–31] for lymphoma, osteosarcoma, and histiocytic sarcoma samples obtained from dogs. From these data, we identified only 4 genomic regions that had highly stable DNA copy numbers among those types of tumors of dogs. In these genomic regions, 9 genes were identified as candidates for use as references for normalization of RT-qPCR assay data. We hypothesized that genes in genomic regions with stable DNA copy numbers would have high expression stability among tumor types.

Results of the present study indicated C_T values for some of the candidate reference genes were highly variable. Results for *SMAD9* indicated an unacceptably high range of C_T values; this gene had one of the lowest expression stabilities in samples of all 3 tumor types evaluated. Genes with a low range of C_T values may be good candidates for use as reference genes because they have consistent expression among tissue types. Results indicated *LAMA1*, *HNF-1*, and *LOC611555* had the lowest ranges of C_T values among tumor types. Of these, *LAMA1* and *HNF-1* had low expression; therefore, the small variation in C_T values for these genes might have been attributable to constitutively low expression. Results of algorithmic analyses also indicated both of those genes had low expression stability in all tumor samples evaluated. High expression was detected for *LOC611555*. This gene also had a low variation in C_T values among tumor types, which indicated it may be useful as a reference gene.

Among candidate reference genes, *LOC611555* had the highest mean expression stability ranking for all 3 tumor types evaluated. Although *LOC611555* was identified as having stable expression among all types of tumors in the present study, use of this gene for normalization of RT-qPCR assay data for other tissue types and for other diseases should be validated because no gene is likely to be optimal for use as a reference in every instance. The *LOC611555* gene is an uncharacterized gene that has a highly conserved sequence in humans, chimpanzees, cows, mice, rats, chickens, and zebrafish.[r] To the authors' knowledge, the function of this gene is unknown. However, given that it had high expression stability among tumor samples obtained from dogs in the present study, it may be useful as a reference for RT-qPCR gene expression data for canine tissue samples.

Expression stability of candidate reference genes identified in the present study was compared with that of 4 genes found to have stable expression in canine tissues by other authors.[23–25] Little information has been published regarding validation of reference genes for normalization of RT-qPCR assay data for canine tissues. Investigators of another study[25] evaluated expression stability of 9 candidate reference genes in prostate, kidney, mammary gland, left ventricle, and liver samples obtained from dogs. Of those genes, *HPRT* had one of the highest expression stabilities. Other authors[23] evaluated expression stability of 11 candidate reference genes in mammary gland samples obtained from healthy dogs and from dogs with disease. In that study,[23] *HPRT*, *RPL32*, *Ubi*, and *ATP5B* were identified as having the highest expression stabilities. Investigators of another study[24] evaluated expression stability of 11 potential reference genes in bone marrow, colon, duodenum, heart, kidney, liver, lung, lymph node, skeletal muscle, pancreas, spleen, and stomach samples obtained from dogs. Results of that study[24] indicated *RPL32* had the highest expression stability in most tissues evaluated and *HPRT* also had high expression stability in bone marrow and lymph node samples. Results of these studies[23–25] indicate *HPRT* and *RPL32* have stable expression among various canine tissues. Results of the present study indicated expression of *HPRT* was stable in histiocytic sarcoma samples and expression of *RPL32* and *Ubi* was stable in osteosarcoma samples, but expression of these genes was not stable in lymphoma samples. Results also indicated expression stability of candidate reference genes identified in the present study was higher than that of genes[23–25] commonly used as references for normalization of RT-qPCR assay data.

Expression stability of genes in tumors of each type obtained from dogs in the present study was ranked by use of 3 computer programs. Program 1 was used to identify suitable reference genes on the basis of a pairwise comparison approach, in which genes with high similarity of expression among RT-qPCR assay samples were assigned a high expression stability rank.[26] Although coregulated genes with expression profiles similar to those identified as having a high expression stability could be assigned a high ranking with program 1, use of programs 2 and 3 avoided this problem.[26,27] Although little annotated information regarding gene function is available for several of the candidate reference genes identified in the present study, determination of expression stability with multiple program algorithms likely yielded more reliable results than would have been obtained by use of only 1 program algorithm. In the present study, ranking of genes on the basis of expression stability varied among program algorithms. However, rankings for the genes with the highest and lowest expression stabilities were similar among the 3 methods. This finding indicated coregulation of gene expression did not likely affect results of the present study.

Use of multiple reference genes for normalization of RT-qPCR assay results is becoming common. We determined expression stability of 13 genes, including 4 genes that are commonly used as reference genes for the 3 tumor types evaluated in the present study. Results obtained by use of program 1 were used to identify several genes that would be useful for multiple-gene normalization of RT-qPCR assay results for each tumor type evaluated. The NF was calculated on the basis of the geometric means of results for genes with the highest expression stabilities. Findings indicated use of data for 2 genes was required for normalization of lymphoma sample RT-qPCR assay results, whereas use of data for 4 genes was required for normalization of histiocytic sarcoma and osteosarcoma sample data.

In the present study, candidate reference genes were identified on the basis of reported[29–31] results of copy number neutrality analyses for > 300 tumor samples obtained from dogs. We identified multiple genes that had higher expression stability among tumor types than did commonly used reference genes. The *LOC611555* gene had the highest expression stability among the 3 tumor types evaluated, and we propose that this gene could be used as a reference to normalize RT-qPCR assay data for lymphoma, histiocytic sarcoma, and osteosarcoma samples obtained from dogs. Normalization factor analysis was used to identify the best combination of reference genes for normalization of RT-qPCR assay data for each tumor type. Results of this study indicated use of commonly used reference genes may not be the best option for normalization of expression data for all types of tumors of dogs, and use of such reference genes may have influenced published findings.

a. geNorm, Center for Medical Genetics, Ghent University, Ghent, Belgium.
b. NormFinder, Molecular Diagnostic Laboratory, Aarhus University Hospital, Aarhus, Denmark.
c. BestKeeper, Institute of Physiology, Center of Life and Food Sciences, Technische Universitat München, Weihenstephan, Germany.
d. CanFam2.0 [database online]. Bethesda, Md: National Center for Biotechnology Information. Available at: www.ncbi.nlm.nih.gov/genome/assembly/237548/. Accessed Sept 30, 2011.
e. UCSC Genome Browser CHORI-82 Canine boxer (*Canis familiaris*) bacterial artificial chromosome library [database online]. Santa Cruz, Calif: UCSC Genome Bioinformatics, University of California-Santa Cruz. Available at: genome.ucsc.edu/cgi-bin/hgGateway. Accessed Sep 30, 2011.
f. RNeasy kit, Qiagen, Valencia, Calif.
g. Turbo DNA free kit, Ambion, Austin, Tex.
h. NanoDrop ND-1000 UV/Visible spectrophotometer, NanoDrop Technologies, Wilmington, Del.
i. Agilent Bioanalyzer 2100, Agilent, Santa Clara, Calif.
j. RefSeq [database online]. Bethesda, Md: National Center for Biotechnology Information. Available at: www.ncbi.nlm.nih.gov/RefSeq/. Accessed Sep 30, 2011.
k. Primer Express, version 3.0, Applied Biosystems, Foster City, Calif.
l. QuantiTect Reverse Transcription Kit, Qiagen, Valencia, Calif.
m. QuantiFast SYBR Green PCR kit, Qiagen, Valencia, Calif.
n. iCycler, Bio-Rad, Hercules, Calif.
o. Fluorescein Passive Reference Dye, Affymetrix, Santa Clara, Calif.
p. GelRed, Biotium, Hayward, Calif.
q. Excel, Microsoft Corp, Redmond, Calif.
r. Gene [database online]. Bethesda, Md: National Center for Biotechnology Information. Available at: www.ncbi.nlm.nih.gov/gene. Accessed Sept 30, 2011.

References

1. Vail DM, MacEwen EG. Spontaneously occurring tumors of companion animals as models for human cancer. *Cancer Invest* 2000;18:781–792.
2. Vail DM, Young KM. Canine lymphoma and lymphoid leukemia. In: Withrow SJ, Vail DM, eds. *Small animal clinical oncology*. 4th ed. Philadelphia: Elsevier Health Sciences, 2007;698–732.
3. Merlo DF, Rossi L, Pellegrino C, et al. Cancer incidence in pet dogs: findings of the Animal Tumor Registry of Genoa, Italy. *J Vet Intern Med* 2008;22:976–984.
4. Priester WA, McKay FW. The occurrence of tumors in domestic animals. *Natl Cancer Inst Monogr* 1980;54:1–210.
5. Bonnett BN, Egenvall A, Olson P, et al. Mortality in insured Swedish dogs: rates and causes of death in various breeds. *Vet Rec* 1997;141:40–44.
6. Lindblad-Toh K, Wade CM, Mikkelsen TS, et al. Genome sequence, comparative analysis and haplotype structure of the domestic dog. *Nature* 2005;438:803–819.
7. Heid CA, Stevens J, Livak KJ, et al. Real time quantitative PCR. *Genome Res* 1996;6:986–994.
8. Higuchi R, Fockler C, Dollinger G, et al. Kinetic PCR analysis: real-time monitoring of DNA amplification reactions. *Biotechnology (N Y)* 1993;11:1026–1030.
9. Bustin SA. Absolute quantification of mRNA using real-time reverse transcription polymerase chain reaction assays. *J Mol Endocrinol* 2000;25:169–193.
10. Thellin O, Zorzi W, Lakaye B, et al. Housekeeping genes as internal standards: use and limits. *J Biotechnol* 1999;75:291–295.
11. Bustin SA. Quantification of mRNA using real-time reverse transcription PCR (RT-PCR): trends and problems. *J Mol Endocrinol* 2002;29:23–39.
12. Dheda K, Huggett JF, Bustin SA, et al. Validation of housekeeping genes for normalizing RNA expression in real-time PCR. *Biotechniques* 2004;37:112–119.
13. Lockwood WW, Chari R, Chi B, et al. Recent advances in array comparative genomic hybridization technologies and their applications in human genetics. *Eur J Hum Genet* 2006;14:139–148.
14. Pinkel D, Albertson DG. Array comparative genomic hybridization and its applications in cancer. *Nat Genet* 2005;37(suppl):S11–S17.
15. Davies JJ, Wilson IM, Lam WL. Array CGH technologies and their applications to cancer genomes. *Chromosome Res* 2005;13:237–248.
16. Fiegler H, Carr P, Douglas EJ, et al. DNA microarrays for comparative genomic hybridization based on DOP-PCR amplification of BAC and PAC clones. *Genes Chromosomes Cancer* 2003;36:361–374.
17. Thomas R, Scott A, Langford CF, et al. Construction of a 2-Mb resolution BAC microarray for CGH analysis of canine tumors. *Genome Res* 2005;15:1831–1837.
18. Thomas R, Duke SE, Karlsson EK, et al. A genome assembly-integrated dog 1 Mb BAC microarray: a cytogenetic resource for canine cancer studies and comparative genomic analysis. *Cytogenet Genome Res* 2008;122:110–121.
19. Ishkanian AS, Malloff CA, Watson SK, et al. A tiling resolution DNA microarray with complete coverage of the human genome. *Nat Genet* 2004;36:299–303.
20. Greshock J, Naylor TL, Margolin A, et al. 1-Mb resolution array-based comparative genomic hybridization using a BAC clone set optimized for cancer gene analysis. *Genome Res* 2004;14:179–187.
21. Vissers LE, de Vries BB, Osoegawa K, et al. Array-based comparative genomic hybridization for the genomewide detection of submicroscopic chromosomal abnormalities. *Am J Hum Genet* 2003;73:1261–1270.
22. Snijders AM, Nowak N, Segraves R, et al. Assembly of microarrays for genome-wide measurement of DNA copy number. *Nat Genet* 2001;29:263–264.
23. Etschmann B, Wilcken B, Stoevesand K, et al. Selection of ref-

erence genes for quantitative real-time PCR analysis in canine mammary tumors using the geNorm algorithm. *Vet Pathol* 2006;43:934–942.
24. Peters IR, Peeters D, Helps CR, et al. Development and application of multiple internal reference (housekeeper) gene assays for accurate normalisation of canine gene expression studies. *Vet Immunol Immunopathol* 2007;117:55–66.
25. Brinkhof B, Spee B, Rothuizen J, et al. Development and evaluation of canine reference genes for accurate quantification of gene expression. *Anal Biochem* 2006;356:36–43.
26. Vandesompele J, De Preter K, Pattyn F, et al. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol* 2002;3:RESEARCH0034.
27. Andersen CL, Jensen JL, Orntoft TF. Normalization of real-time quantitative reverse transcription-PCR data: a model-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets. *Cancer Res* 2004;64:5245–5250.
28. Pfaffl MW, Tichopad A, Prgomet C, et al. Determination of stable housekeeping genes, differentially regulated target genes and sample integrity: BestKeeper—Excel-based tool using pairwise correlations. *Biotechnol Lett* 2004;26:509–515.
29. Thomas R, Seiser EL, Motsinger-Reif AA, et al. Refining tumor-associated aneuploidy through 'genomic recoding' of recurrent DNA copy number aberrations in 150 canine non-Hodgkin's lymphomas. *Leuk Lymphoma* 2011;52:1321–1335.
30. Angstadt AY, Motsinger-Reif AA, Thomas R, et al. Characterization of canine osteosarcoma by array comparative genomic hybridization and RT-qPCR: signatures of genomic imbalance in canine osteosarcoma parallel the human counterpart. *Genes Chromosomes Cancer* 2011;50:859–874.
31. Hedan B, Thomas R, Motsinger-Reif AA, et al. Molecular cytogenetic characterization of canine histiocytic sarcoma: a spontaneous model for human histiocytic cancer identifies deletion of tumor suppressor genes and highlights influence of genetic background on tumor behavior. *BMC Cancer* 2011;11:201.
32. Rozen S, Skaletsky HJ. Primer3 on the WWW for general users and for biologist programmers. In: Krawetz S, Misener S, eds. *Bioinformatics methods and protocols: methods in molecular biology*. Totowa, NJ: Humana Press, 2000;365–386.
33. Maccoux LJ, Clements DN, Salway F, et al. Identification of new reference genes for the normalisation of canine osteoarthritic joint tissue transcripts from microarray data. *BMC Mol Biol* 2007;8:62.