Ann. Anim. Sci., Vol. 19, No. 3 (2019) 619–632 *DOI: 10.2478/aoas-2019-0027*

sciendo

ANALYSIS OF MITOCHONDRIAL GENOME FROM LABRADOR (*CANIS LUPUS FAMILIARIS*) WITH MAMMARY GLAND TUMOUR REVEALS NOVEL MUTATIONS AND POLYMORPHISMS

Krzysztof Kowal[1], Brygida Ślaska[1]♦, Adam Bownik[1], Beata Horecka[1], Jan Gawor[2], Anna Śmiech[3], Angelika Tkaczyk[1]

[1]Institute of Biological Bases of Animal Production, Faculty of Biology, Animal Sciences and Bioeconomy, University of Life Sciences in Lublin, Akademicka 13, 20-950 Lublin, Poland
[2]Laboratory of DNA Sequencing and Oligonucleotide Synthesis, Institute of Biochemistry and Biophysics, Polish Academy of Sciences, Pawińskiego 5a, 02-106 Warsaw, Poland
[3]Sub-Department of Clinical Diagnostic and Veterinary Dermatology, Faculty of Veterinary Medicine, University of Life Sciences in Lublin, Głęboka 30, 20-612, Lublin, Poland
♦Corresponding author: brygida.slaska@up.lublin.pl

Abstract

The aim of the study was to find associations between the process of neoplastic transformation and mtDNA mutations/polymorphisms, i.e. factors with potential prognostic significance, and to determine their impact on the biochemical properties, as well as structural, and functional properties of proteins. Blood and neoplastic tissue samples were collected from a 9-year-old Labrador dog with a diagnosed malignant mammary tumour. Next-generation genome sequencing (NGS) of the entire mitochondrial genome was performed using Illumina technology, and bioinformatics analyses were carried out. This is the first report demonstrating the application of NGS in the analysis of the canine mtDNA genome in neoplastic disease. The proposed strategy is innovative and promising. For the first time in the literature, the sequence of 29 genes was analysed to determine their association with the prevalence of tumour. In total, 32 polymorphic loci and 15 mutations were identified. For the first time, as many as 24 polymorphisms and all the mutations have been described to be associated with the neoplastic process in dogs. Most polymorphisms/mutations were found in the D-loop (31% of the polymorphisms and 93% of the mutations) and the *COX1* gene sequence (16% of the polymorphisms). Blood or cancer heteroplasmy was noted in 93% of the mutations. Four of the 18 polymorphisms detected in the protein-coding genes were non-synonymous polymorphisms that have not been described in the literature so far (m.T7593C in *COX2*, m.G8807A in *COX3*, m.A9911G in *ND4L*, and m.T13299A in *ND5*) but resulted in changes in amino acids in proteins. These mutations and polymorphisms can affect mitochondrial functions and may be a result of cell adaptation to the changes in the environment occurring during carcinogenesis. The replacement of "wild type" mtDNA by a mutated molecule may be an important phenomenon accompanying carcinogenesis.

Key words: tumour, dog, mtDNA, mutations, NGS

The role of mitochondria in the malignant process in humans was determined already over 70 years ago when, in 1956, Otto Warburg described the glycolysis phenomenon occurring in cancer cells despite the presence of oxygen. The high level of glycolysis in the presence of oxygen (called the Warburg effect) results from genetic or epigenetic changes and is characteristic of most cancers. Metabolic changes such as elevated levels of glycolysis and increased glucose consumption play an important role in the process of malignant transformation, increase the aggressiveness and invasiveness of tumours, and act as antiapoptotic factors (Grzybowska-Szatkowska and Slaska, 2012).

Mitochondrial DNA (mtDNA) is characterised by a ten-fold higher mutagenicity rate than its nuclear counterpart. It is associated with e.g. a high concentration of free oxygen radicals generated in the oxidative phosphorylation process. DNA strands that are not covered with histone proteins are especially susceptible to the action of reactive oxygen species, which also contribute to an increase in the frequency of mutations. A relationship has been found between mutations in human mtDNA in the tumour initiation and progression phases supporting the genetic instability of the cell and disrupting its repair functions (Bai et al., 2009; Grzybowska-Szatkowska and Slaska, 2012; Ślaska et al., 2013).

The length of the mtDNA molecule of *Canis lupus familiaris* was found to be 16,728 bp, having a 1270 bp-long control region including 30 tandem repeats of a 10-nt motif between conserved sequence block (CSB) I and II. 37 genes are encoded in the mitochondrial genome (Kim et al., 1998). 28 CG nucleotide-rich genes are encoded on the heavy strand (H), including 10 genes encoding respiratory chain proteins, 2 ATPase genes, 14 tRNA genes, and 2 rRNA genes. The other tRNA genes and the *NADH6* gene are encoded on the H-strand so that the L-strand gives the sense reading frames (Kim et al., 1998).

Although the mitochondrial genome sequence was described at the end of the 20th century, the association of mtDNA mutations with neoplastic transformation in dogs was investigated much later. However, it was noted that the aetiology of many diseases, including tumours, is similar at the molecular level in both dogs and humans (Ślaska et al., 2013). Investigations of the association of mitochondrial DNA mutations with the development of human cancers are focused on both coding and non-coding gene regions (Grzybowska-Szatkowska and Slaska, 2012). Somatic mtDNA mutations have been reported to be associated with neoplastic transformation in dogs as well (Slaska et al., 2014, 2015; Ślaska et al., 2014, 2016; Śmiech et al., 2016, 2019; Surdyka and Slaska, 2017 a, b), but most mutations have been described in the non-coding region of the mtDNA D-loop. The research of coding genes in canine mtDNA has been limited to only a few genes (Slaska et al., 2015; Ślaska et al., 2016; Surdyka and Slaska, 2017 b; Śmiech et al., 2019). A vast majority of mtDNA genes were not analysed in those studies. To date, the world literature does not provide information about mutations and polymorphisms in the entire mitochondrial genome associated with neoplastic transformation in dogs. Given the clonal nature of mtDNA and the high copy number, mtDNA mutations can serve as a new effective molecular biomarker for detection of canine tumours. It is plausible that determination of the mtDNA defect may be useful in the diagnosis and detection of tumours.

The aim of the study was to determine the association of the neoplastic transformation process with mtDNA mutations/polymorphisms, i.e. factors with potential prognostic importance, and to identify their impact on the biochemical properties, structure, and function of proteins. Another objective was to detect regions in the entire mitochondrial genome that are particularly susceptible to the occurrence of mutations and polymorphisms in canine mammary tumours.

Material and methods

Animals

The analyses were carried out on tissue samples collected from a 9-year-old Labrador dog with diagnosed mammary tumour (tubulopapillary carcinoma, G2). The dog received neither hormone therapy nor chemotherapy. DNA extracted from postoperative cancer tissue and blood of the examined dog was analysed.

Laboratory procedures

DNA was isolated with the use of an automated nucleic acid extraction system – QIACube (Qiagen, Hilden, Germany). The DNeasy Blood & Tissue Kit (Qiagen) was used for DNA extraction from mammary gland tumour tissue and blood. The DNA sample was assessed qualitatively by electrophoretic separation in agarose gel and qualitatively (BioPhotometer spectrophotometer; Eppendorf, Hamburg, Germany). The mammary gland tumour tissue sample was placed in a sterile container. Blood was sampled into sterile test-glasses with the K_2EDTA anticoagulant (dipotassium ethylenediaminetetraacetic acid) (Medlab, Raszyn, Poland). In order to determine the histological types of the analysed mammary gland tumour tissue, a section was sampled from surgically removed tumour for histopathological analysis. The mammary gland tumour tissue was fixed in buffered formalin, embedded in paraffin blocks, and routinely stained with H&E and toluidine blue. Microscopic assessment of the tumour was performed. The malignancy degree of the mammary gland tumour was assessed using the 3-grade scale of malignancy, i.e. a sum of point values assigned to histomorphological traits according to Goldschmidt et al. (2011).

Mitochondrial DNA was selectively amplified on the total genomic DNA using two pairs of primers: F1418 and R11041, ~9,5kb PCR product, and 9190F and R2382, ~9,8kb PCR product (Imes et al., 2012). These two long-range PCR amplicons fully covered the mitochondrial genome sequence. PCR reactions were carried out using KAPA HiFi PCR Kit reagents (KAPA Biosystems, Wilmington, USA). After amplification, the PCR products were analysed in agarose gel and purified using Ampure XP magnetic beads (Beckman Coulter, Brea, USA). The quantity of DNA was measured using a Qubit 2.0 fluorimeter (Thermo Fisher Scientific, Waltham, USA). Both PCR products were diluted to the desired concentration and pooled in an equimolar ratio. Such a DNA mixture was next used in library preparation.

Mitochondrial DNA (mtDNA) was sequenced in the DNA Sequencing and Oligonucloetide Synthesis Laboratory (oligo.pl) at the Institute of Biochemistry and Biophysics Polish Academy of Science. Approximately 1 µg of the PCR DNA template mix was mechanically sheared by nebulization and Illumina shotgun library was constructed using a TruSeq PCR-Free Kit (Illumina, San Diego, CA) following manufacturer's instructions. The library sample was sequenced on an Illumina MiSeq sequencer (Illumina, San Diego, CA) using a 600-cycle kit (v3) in paired-end mode. The sequence reads were filtered by quality using the FastX toolkit (http://hannonlab.cshl.edu/fastx_toolkit/) and assembled into contigs with the use of the Newbler v.3.0 GS De Novo Assembler (Roche, USA). Contig alignments were investigated using Seqman software from Lasergene package (DNAStar, USA). The remaining gaps were closed by PCR and confirmed by Sanger capillary sequencing.

Bioinformatics analysis

The nucleotide sequences obtained were subjected to bioinformatics analyses (Altschul et al., 1997; Huang and Madan, 1999) in order to determine mutation and polymorphic sites within the analysed mtDNA genome in the neoplasm tissue (DNA Baser Sequence Assembler v3.2 (2012), Unipro UGENE (v. 1.26.3.), and MEGA (v. 7.0.26)). The nucleotide sequence in the neoplasm tissue and in the blood was compared to the reference sequence (GenBank accession NC_002008; Kim et al., 1998). The analysis of mutations in mt-tRNA genes and their influence on the tRNA structure was conducted in tRNAscan-SE 2.0 (Lowe and Chan, 2016).

The probability of deleterious mutations, i.e. a functional effect of the non-synonymous protein-coding SNP, was determined using the Panther Classification System (Tang and Thomas, 2016). It predicts disease-causing genetic variants using position-specific evolutionary preservation. The ExPASy Server (Gasteiger et al., 2005) was used to characterise such physicochemical parameters as the theoretical isoelectric point (pI), instability index, aliphatic index, and grand average hydropathy (GRAVY). SOPMA was used for calculating the secondary structural features of the antioxidant protein sequences considered. Trans Membrane prediction using the Hidden Markov Model (TMHMM) was used for predicting transmembrane helices based on the Hidden Markov Model (Combet et al., 2000).

The investigations consisted in analyses of tubulo-papillary carcinoma, as this is one of the most prevalent canine mammary tumours diagnosed in Poland (Łopuszyński et al., 2010; Szczubiał and Łopuszyński, 2011).

Results

The Illumina sequencing yielded 19 182 raw paired reads for the tumour sample and 11 862 raw reads for the blood sample (Table 1). After quality trimming following sequencing adaptor removal, 16 164 paired reads and 9 671 988 nucleotides of sequence data remained for the tumour sample. In the case of the blood sample, 10 273 paired reads and 6 152 592 nucleotides of the sequencing data passed the

quality filters (Table 1). Mitochondrial genome assembly using the Newbler program generated a single circular contig with the size of 16642 bp in both cases. 580X sequencing coverage was generated for the tumour sample and 365X genome coverage was obtained for the blood sample sequentially (Table 1).

Table 1. Summary of mitochondrial DNA sequencing results

Sample	Reads count		Sequencing data yield (nt)	mtDNA genome coverage
	Raw reads	Filtered reads		
Blood	11 862	10 273	6 152 592	580x
Tumor	19 182	16 164	9 671 988	365x

Polymorphisms and/or mutations were detected in 10 of the 13 protein-coding mtDNA genes in the Labrador dog diagnosed with tubulopapillary carcinoma. They were also noted in the *16S rRNA* gene and in two genes encoding mt-tRNA (*TRNL1* (coding tRNA-Leu1), *TRNR* (coding tRNA-Arg)). The highest number of mutations/polymorphisms was recorded in the non-coding region of the D-loop (Tables 2 and 3). There were no mutations and polymorphisms in three protein-coding genes (*ATP8*, *NADH3*, *NADH6*), the *12S rRNA*-encoding gene, and 20 tRNA-encoding genes. A majority of the mutations and polymorphisms were substitutions, while insertions were detected in three mtDNA positions (Tables 2 and 3).

The presence of polymorphisms in relation to the canine reference sequence was detected in 32 positions of mitochondrial DNA (Table 2), and the highest number was noted in the D-loop (10 polymorphic sites) and in the *COX1* gene sequence (5 polymorphisms). Four of the 18 polymorphisms present in the protein-coding genes were non-synonymous polymorphisms resulting in alterations in the protein amino acids (Table 2). Polymorphism m.T7593C in the *COX2* gene resulted in substitution of methionine with threonine (p.M187T) in the COX2 protein. The substitution of cysteine with tyrosine (p.C55Y) in the COX3 protein was caused by the m.G8807A polymorphism. In turn, the m.A9911G polymorphism led to substitution of methionine with valine (p.M1V) in the ND4L protein. The m. T13299A polymorphism contributed to substitution of serine with threonine (p.S508T) in the ND5 protein (Table 2). There were also three insertion-type length polymorphisms, i.e. two in genes encoding mt-tRNA (*TRNL1* (coding tRNA-Leu1) (UUR) – m.2678_2679insG and *TRNR* (coding tRNA-Arg) – m.9910_9911insAT), and one in the D-loop (m.16660_16661insCC). The changes were located in the D-arm and loop in the *TRNL1* gene and in the acceptor stem region in the *TRNR* gene.

Somatic mutations were detected in 15 positions of the mitochondrial genome, with 14 identified as transition mutations in the hypervariable mtDNA region (Table 3). Additionally, there was a non-synonymous mutation in position m.11028 in the *ND4* gene sequence (Table 3).

Nearly all the identified mutations were heteroplasmic. Blood or cancer heteroplasmy was noted in 14 of the 15 mutations. Cancer heteroplasmy was identified in seven positions in the hypervariable mtDNA region (m.16168A/G, m.16178A/G, m.16188A/G, m.16198A/G, m.16228A/G, m.16238A/G, m.16358A/G), whereas blood heteroplasmy was detected in six positions of the D-loop (m.16278A/G,

m.16288A/G, m.16328A/G, m.16338A/G, m.16368A/G, m.16388A/G) and in gene *ND4* (m.11028C/T) (Table 3).

Table 2. Polymorphisms in the mitochondrial genome in the reference sequence, blood, and tumour tissue of the Labrador dog with tubulopapillary carcinoma

Gene	Position in reference sequence (1..16727)	Reference sequence NC_002008	Sequences in blood	Sequences in tumour cells	Amino acid change	Position of the changed nucleotide in the codon
16S rRNA	1091..2670	m.1351A	m.A1351G	m.A1351G	–	–
TRNL1	2671..2744	m.2678_2679	m.2678.1G	m.2678.1G	–	–
		m.2683G	m.G2683A	m.G2683A	–	–
ND1	2747..3703	m.2962C	m.C2962T	m.C2962T	p.I72I	ATC→ATT
		m.3196T*	m.T3196C	m.T3196C	p.L150L	CTT→CTC
ND2	3914..4957	m.4940T	m.T4940C	m.T4940C	p.L343L	CTA→TTA
COX1	5349..6893	m.5367C	m.C5367T	m.C5367T	p.L7L	CTG→TTG
		m.5444T	m.T5444C	m.T5444C	p.A32A	GCT→GCC
		m.6065A	m.A6065G	m.A6065G	p.G239G	GGA→GGG
		m.6401C	m.C6401T	m.C6401T	p.G351G	GGC→GGT
		m.6554T	m.T6554C	m.T6554C	p.G402G	GGT→GGC
COX2	7034..7717	m.7593T	m.T7593C	m.T7593C	p.M187T	ATA→ACA
ATP6	7964..8644	m.8281T*	m.T8281C	m.T8281C	p.I106I	ATT→ATC
		m.8368C	m.C8368T	m.C8368T	p.L135L	CTC→CTT
COX3	8644..9427	m.8807G	m.G8807A	m.G8807A	p.C55Y	TGC→TAC
TRNR	9842..9910	m.9910_9911	m.9910.2AT	m.9910.2AT	–	–
ND4L	9911..10207	m.9911A	m.A9911G	m.A9911G	p.M1V	ATG→GTG
ND4	10201..11578	m.10611A	m.A10611T	m.A10611T	p.G137G	GGA→GGT
		m.10992G*	m.G10992A	m.G10992A	p.L264L	TTG→TTA
ND5	11778..13598	m.13299T	m.T13299A	m.T13299A	p.S508T	TCA→ACA
CYB	14183..15322	m.14977T*	m.T14977C	m.T14977C	p.P265P	CCT→CCC
		m.15214G	m.G15214A	m.G15214A	p.E344E	GAG→GAA
D-loop	15458..16727	m.15620T*	m.T15620C	m.T15620C	–	–
		m.15627A*	m.A15627G	m.A15627G	–	–
		m.15639T*	m.T15639A	m.T15639A	–	–
		m.15814C*	m.C15814T	m.C15814T	–	–
		m.16148A	m.A16148G	m.A16148G	–	–
		m.16158A	m.A16158A/G	m.A16158A/G	–	–
		m.16398A	m.A16398G	m.A16398G	–	–
		m.16418A	m.A16418G	m.A16418G	–	–
		m.16660_16661	m.16660.2CC	m.16660.2CC	–	–
		m.16672C	m.C16672T	m.C16672T	–	–

*mtDNA position described in literature reports of canine tumours.

Table 3. Mutations in the mitochondrial genome between blood and tumour tissue of the Labrador dog with tubulopapillary carcinoma

Gene/Region	Reference sequence NC_002008 (1..16727)	Sequences in blood	Sequences in tumour cells	Amino acid change	Position of the changed nucleotide in the codon
ND4	m.11028T	m.11028C/T*	m.11028C	p.C276C	TGT→TGC
D-loop	m.16168A	m.16168A	m.16168A/G*	–	–
	m.16178A	m.16178A	m.16178A/G*	–	–
	m.16188G	m.16188G	m.16188A/G*	–	–
	m.16198G	m.16198A	m.16198A/G*	–	–
	m.16218G	m.16218G	m.16218A	–	–
	m.16228G	m.16228G	m.16228A/G*	–	–
	m.16238G	m.16238A	m.16238A/G*	–	–
	m.16278G	m.16278A/G*	m.16278G	–	–
	m.16288A	m.16288A/G*	m.16288A	–	–
	m.16328G	m.16328A/G*	m.16328G	–	–
	m.16338G	m.16338A/G*	m.16338A	–	–
	m.16358A	m.16358A	m.16358A/G*	–	–
	m.16368G	m.16368A/G*	m.16368A	–	–
	m.16388A	m.16388A/G*	m.16388G	–	–

*Heteroplasmy.

In the region from 16130 to 16430 bp of the D-loop, there is a 10-nucleotide tandem repeat motif (5`-GTACACGT(A/G)C-3`). In the case of tubulopapillary carcinoma, this motif exhibits greater variation than that of the reference sequence (Table 4). Among the 30 tandem repeats, 12 were found to retain the 5'-GTACACGTAC-'3 motif in the blood and 11 retained the motif in the tumour tissue. The same numbers of 5'-GTACACGTGC-'3 repeats were observed in both the blood and tumour, and these numbers were lower than in the reference sequence. In the case of tubulopapillary carcinoma, there was also an increased number of motifs, in which the 5'-GTACACGTA/GC-'3 read indicates heteroplasmy.

Table 4. Variation in the number of tandem repeat motifs between the reference sequence, blood, and tumour tissue sampled from the canine mammary tumour

	Number of 10-nucleotide motifs in the VNTR region		
	5'-GTACACGTAC-'3	5'-GTACACGTGC-'3	5'-GTACACGTA/GC-'3
Reference sequence	16/30	14/30	0/30
Sequences in blood	12/30	11/30	7/30
Sequences in tumour cells	11/30	11/30	8/30

The analyses carried out to determine the deleterious effect of the polymorphisms, i.e. a functional effect of the non-synonymous protein-coding SNP determined using

the Panther Classification System in the case of the substitution of methionine with valine (p.M1V) in the ND4L protein indicate that a substitution of a single amino acid in the protein has a probably damaging impact on its function. The analyses carried out to determine the deleterious effect of substitutions p.M187T in the COX2 protein, p.C55Y in the COX3 protein, and p.S508T in the ND5 protein indicated that the substitutions were more likely not to disrupt protein function.

In the case of the amino acid substitution p.M1V in the ND4L protein, the instability index value indicates a transition in the stability of the protein from unstable in the reference sequence to stable (Table 5) at the m.A9911G mutation in the *ND4L* gene sequence (Table 2). No significant differences were noted in the case of the other protein properties in the non-synonymous protein-coding SNP of the *COX2*, *COX3*, and *ND5* genes of the dog with tubulopapillary carcinoma (Table 5). Secondary structure elements are presented in Table 6. The results revealed that the random coil dominated in the secondary structure of the COX2 protein (p.M187T), followed by the alpha helix, extended strand, and beta turns. The alpha helix dominated in the secondary structure of the COX3 (p.C55Y), ND4L (p.M1V), and ND5 (p.S508T) proteins, followed by the random coil, extended strand, and beta turns. There was only a slight decrease in the percentage proportion of the compact alpha helix structure in the case of the p.1M1V substitution (73.47%) relative to normal protein (83.67%) and thus an increase in the percentage of the random coil and extended strand structures (Table 6).

Table 5. Comparison of protein properties in non-synonymous protein-coding SNP in the canine tumour calculated by the ExPASy Server

Protein	Amino acid change	Theoretical pI (Isoelectric point)	Aliphatic index	Instability index	Grand average of hydropathicity (GRAVY)
COX2	p.M187T	4.62	109.91	36.97 (stable protein)	0.285
	Normal	4.62	109.91	36.64 (stable protein)	0.296
COX3	p.C55Y	6.44	93.37	20.87 (stable protein)	0.444
	Normal	6.44	93.37	19.67 (stable protein)	0.459
ND4L	p.M1V	5.27	143.16	38.74 (stable protein)	1.326
	Normal	5.27	140.20	43.22 (unstable protein)	1.302
ND5	p.S508T	9.28	109.95	31.07 (stable protein)	0.591
	Normal	9.28	109.95	31.51 (stable protein)	0.591

The results indicate that the non-synonymous mutations resulting in changes in the amino acids (p.M187T, p.C55Y, p.M1V, p.S508T) in the tubulopapillary carcinoma tissue did not influence the number of transmembrane sequences (Table 7). The values of the expected number of amino acids in transmembrane helices (Exp number of AAs in TMHs) of the COX3, ND4L, and ND5 proteins, which substantially exceeded the threshold value (18), suggest that they are very likely to be transmembrane proteins or serve as signal peptides. The expected number of amino acids in transmembrane helices in the first 60 amino acids of the protein (Exp number, first

60 AAs) was comparable to that for normal protein and COX2, COX3, ND4L, and ND5 proteins with amino acid changes (Table 7).

Table 6. Secondary structure elements calculated by SOPMA

Secondary structure	COX2		COX3		ND4L		ND5	
	p.M187T	normal	p.C55Y	normal	p.M1V	normal	p.S508T	normal
Alpha helix (%)	29.96	31.72	50.19	50.96	73.47	83.67	55.12	55.61
310 helix (%)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Pi helix (%)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Beta bridge (%)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Extended strand (%)	23.35	24.23	17.24	15.71	9.18	7.14	13.53	13.04
Beta turn (%)	5.73	3.52	5.75	6.90	2.04	2.04	3.63	3.14
Bend region (%)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Random coil (%)	40.97	40.53	26.82	26.44	15.31	7.14	27.72	28.22
Ambiguous states (%)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Other states (%)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Table 7. Prediction of transmembrane helices in genes using the Hidden Markov Model (TMHMM)

	Protein name/ Amino acid change	Length (aa)	Transmembrane sequences	Exp number of AAs in TMHs	Exp number, first 60 AAs
COX2	p.M187T	183	1	22.55	22.34
	normal	183	1	22.59	22.34
COX3	p.C55Y	217	4	102.82	21.78
	normal	217	4	102.82	21.78
ND4L	p.M1V	98	3	62.86	41.82
	normal	98	3	62.77	41.82
ND5	p.S508T	562	12	270.60	20.94
	normal	562	12	270.62	20.94

Discussion

Investigations of the association of mitochondrial DNA with neoplastic transformation in dogs have been conducted for only a few years. They are partly focused on the non-coding region of the mitochondrial genome, i.e. the D-loop, in which numerous mutations and polymorphisms associated with canine tumours have been detected (Ślaska et al., 2014; Slaska et al., 2014; Śmiech et al., 2016; Surdyka and Slaska 2017 a). The present study demonstrated 31% of all described polymorphisms and over 93% of mutations in the D-loop (mtDNA nucleotide position 15458..16727) in the dog with mammary tumour (Tables 2 and 3). This may indicate that, despite its non-coding character, the D-loop plays an important role in the neoplastic transfor-

mation process. This thesis is confirmed by the results obtained in this study (Tables 2 and 3) and literature data (Ślaska et al., 2014; Slaska et al., 2014; Śmiech et al., 2016; Surdyka and Slaska, 2017 a).

Four of the ten polymorphisms in the D-loop (Table 2) have already been shown in the literature to be associated with the neoplastic process in dogs. They were detected in mtDNA positions m.15620, m.15627, m.15639 (Slaska et al., 2014 ; Ślaska et al., 2014; Surdyka and Slaska, 2017 a), and m.15814 (Śmiech et al., 2016; Surdyka and Slaska, 2017 a) (Table 2). This may imply a predilection of dogs with tumour diseases to carry a D-loop defect. The other polymorphisms in positions m.16148, m.16158, m.16398, m.16418, m.16660_16661, and m.16672 (Table 2) have been identified here for the first time, probably due to the fact that the D-loop region starting from position 16148 bp of mtDNA has not been investigated in relation to canine tumours so far.

None of the mutations detected in tubulopapillary carcinoma (Table 3) has been described in the literature as a mutation associated with the neoplastic process in dogs. Considering the non-coding mtDNA region, this may be associated with the fact that the D-loop region has not been investigated in the literature. The present study describes for the first time a non-synonymous mutation in the *ND4* gene in the form of blood heteroplasmy (m.11028C/T) (Table 3). Mutations in the *ND4* gene in tumours diagnosed in German Shepherd dogs were presented by Ślaska et al. (2016); however, the mutation at position m.11028 was not reported.

Blood heteroplasmy was noted in six positions in the D-loop mutations, whereas tumour heteroplasmy was detected in seven D-loop positions (Table 3). It should be emphasised that blood or tumour heteroplasmy was detected in almost 93% of the mutations identified in the D-loop.

As in the study conducted by Surdyka and Slaska (2017 a), we observed length heteroplasmy in the Polyc-polyT-Polyc stretch in position m.16660.2CC (Table 2). It is possible that the accumulation and conservation of numerous SNPs in the genome have an impact on the progression and/or development of canine mammary cancer, since SNP in the D-loop region may induce ROS generation and contribute to neoplastic transformation, as shown by other authors (Brandon et al., 2006). Length heteroplasmy is a common phenomenon in the mitochondrial genome. This may indicate a potential use of the heteroplasmy phenomenon in the diagnostics of canine tumours.

Noteworthy, the mutations in the D-loop were present only in the 220 bp fragment (position 16168-16388 bp of mtDNA), which accounts for merely 17% of its length (Table 3). This analysed fragment is located entirely within the 16130-16430 bp region of the D-loop comprising motif 5'-GTACACGTA/GC-'3 in the hypervariable VNTR region (Kim et al., 1998; Webb and Allard, 2009), in which there were differences between the analysed sequences present in the tumour, blood, and reference sequence associated with the frequency of the motif variants and heteroplasmy (Table 4). Heteroplasmy can be hereditary when two versions of the mitochondrial genome are inherited from the mother (Nisztuk-Pacek et al., 2019). It can also be associated with the emergence of *de novo* mutations, e.g. in tumour tissue (Slaska et al., 2014; Ślaska et al., 2014; Śmiech et al., 2016; Surdyka and Slaska,

2017 a). As shown by the available literature, this region has not been analysed to date with regard to the association of mutations with the neoplastic process in dogs. Yet, this issue should be further investigated with great attention, given its potential prognostic significance. This assumption is reasonable, since the mtDNA D-loop is responsible for the process of replication, transcription, and organisation of the mitochondrial genome. The available data suggest that mutations in this region may impair the function of the electron transport chain, thereby leading to generation of increased levels of reactive oxygen species, which damage the DNA structure. Additionally, excess levels of reactive oxygen species can lead to nuclear DNA damage and, consequently, tumour development (Brandon et al., 2006; Dement et al., 2007; Grzybowska-Szatkowska and Slaska, 2012).

The present study demonstrated yet undescribed numerous polymorphisms in the mtDNA coding regions associated with the neoplastic process in dogs. Among the 13 mtDNA coding genes, the following have been described to date: *ND1*, *ND2*, *ND4*, *COX1, COX2, COX3*, *ATP6*, and *CYB* (Slaska et al., 2015; Ślaska et al., 2016; Surdyka and Slaska, 2017 b; Śmiech et al., 2019). However, only four canine tumour-related polymorphisms of the 22 polymorphisms detected in the present investigations (Table 2) were described in the studies mentioned above. These included polymorphisms in the *ND1* gene in position m.3196 of mtDNA (Slaska et al., 2015), in position m.8281 of the *ATP6* gene (Surdyka and Slaska, 2017 b), in position m.10992 of the *ND4* gene (Ślaska et al., 2016), and in position m.14977 in the *CYB* gene (Śmiech et al., 2019). This indicates that 18 new polymorphisms related to the neoplastic process in dogs have been detected in the present study (Table 2).

Noteworthy are the four yet undescribed non-synonymous polymorphisms (m.T7593C in *COX2*, m.G8807A in *COX3*, m.A9911G in *ND4L*, and m.T13299A in *ND5*) detected in the protein-coding genes (Table 2), due to their potential adverse effects on the function of proteins. The analysis of the main protein parameters (Tables 5–7), except for the ND4L protein, did not reveal significant differences in the values of parameters between the analysed sequences, which indicates that the substitutions are more likely not to disrupt protein function. It is probable that although there was no significant effect of the amino acid changes in positions p.M187T, p.C55Y, and p.S508T induced by mutations on genes *COX2*, *COX3* and *ND5*, respectively, on the protein function, the total effect of the accompanying polymorphisms may have contributed to neoplastic transformation. Therefore, genes *COX2*, *COX3*, *NDL4*, and *ND5* should be subjected to further analyses.

The change in the ND4L protein (p.M1V) associated with the non-synonymous polymorphism in the *ND4L* gene and described in the literature for the first time is puzzling. *In silico* analyses have demonstrated that the difference in one position of the amino acid sequence compared with the reference sequence has a positive effect on protein stability, i.e. the change transforms an unstable protein before the mutation into a stable one after the mutation (Table 5). This can be verified by the theory proposed by Guruprasad et al. (1990), who argue that protein instability is caused by the presence of certain dipeptides in the protein structure. This study showed a mutation that resulted in a change in the amino acid sequence p.M1V. Another amino acid in the peptide sequence is serine. The instability value is 44.94 for di-

peptide MS and 1.0 for VS (Guruprasad et al., 1990), which explains the increase in the stability of the ND4L structure. In the case of neoplastic transformation, the mutation results in transition into the stable structure of the ND4L protein. However, the analyses carried out to determine the functional effect of the non-synonymous protein-coding SNP, in the case of substitution of methionine with valine (p.M1V) in the ND4L protein, indicate probability that a substitution of a single amino acid in the protein has a probably detrimental impact on its function.

The sequence of 29 genes was analysed for the first time to assess their association with the occurrence of tumours. 13 genes out of this number exhibited polymorphisms (Table 2). Additionally, many polymorphisms in coding genes were detected and ascribed a role in the neoplastic process in dogs for the first time (Table 2). Their greatest number, i.e. 5 yet undescribed polymorphisms, was detected in the *COX1* gene (Table 2). The literature provides data on two synonymous polymorphisms in the *COX1* gene associated with canine tumours (Slaska et al., 2015); yet, none of them has been found in this study. Given the high number of the tumour-associated polymorphisms of the *COX1* gene presented in this study and in the literature, continuation of the research of the *COX1* gene in terms of typing tumour markers in dogs is advisable.

To date, mutations and polymorphisms in mt-RNA genes have not been analysed in association with canine tumours. The present study demonstrated two insertion-type length polymorphisms in two mt-tRNA-coding genes (*TRNL1* (coding tRNA-Leu1) and *TRNR* (coding tRNA-Arg)) out of 22 mt-tRNA-coding genes (Table 2). *In silico* analyses indicate changes in the D-arm and loop of the *TRNL1* gene and in the acceptor stem region of the *TRNR* gene. The results of studies conducted by McFarland et al. (2004) have shown that pathogenic mutations of mt-RNA genes in humans are more often located at conservative sites and stem structures and tend to disrupt Watson–Crick (WC) nucleotide pairing in stems. In the analysed material, there was an insertion within the terminal fragment of the *TRNR* gene, which may exert a negative effect on the formation of the aminoacyl-tRNA synthase complex. Polymorphisms in the structure of the tRNA-Leu molecule probably do not impair its function (Panwar and Raghava, 2014; Lowe and Chan, 2016). It is unlikely that genes for tRNA are 'hot spots' for canine mammary tumours; however, it cannot be excluded that, through their impact on the secondary and tertiary tRNA structure, polymorphisms may cause mitochondrial dysfunction.

In this study, 32 polymorphic loci and 15 mutations were identified. As many as 24 polymorphisms and all the mutations have been described to be associated with neoplastic transformation in dogs. The presence of a large number of polymorphisms and heteroplasmic mutations in the D-loop may indicate that, despite its non-coding nature, the D-loop plays a vital role in the neoplastic transformation process. Mutations in tumours may indicate genetic instability of mtDNA, which may also be involved in carcinogenesis.

Ethics approval

The study was approved by the II Local Ethical Commission for animal experiments in Lublin, Poland (resolution number 6/2013).

References

Altschul S.F., Madden T.L., Schaffer A.A., Zhang J., Zhang Z., Miller W., Lipman D.J. (1997). Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. Nucleic Acids Res., 25: 3389–3402.

Bai Y., Lu J., Sharma K.L. (2009). Implications of mitochondrial DNA mutations and mitochondrial dysfunction in tumorigenesis. Cell Res., 19: 802–815.

Brandon M., Baldi P.A., Wallace D.C. (2006). Mitochondrial mutations in cancer. Oncogene, 25: 4647–4662.

Combet C., Blanchet C., Geourjon C., Deleage G. (2000). NPS@: network protein sequence analysis. Trends Biochem. Sci., 25: 147–150.

Dement G.A., Maloney S.C., Reeves R. (2007). Nuclear HMGA1 nonhistone chromatin proteins directly influence mitochondrial transcription, maintenance, and function. Exp. Cell Res., 313: 77–87.

Gasteiger E., Hoogland C., Gattiker A., Wilkins M.R., Appel R.D., Bairoch A. (2005). Protein identification and analysis tools on the ExPASy server. In: The proteomics protocols handbook. Humana Press, pp. 571–607.

Goldschmidt M., Peña L., Rasotto R., Zappulli V. (2011). Classification and grading of canine mammary tumors. Vet. Pathol., 48: 117–131.

Grzybowska-Szatkowska L., Slaska B. (2012). Mitochondrial DNA and carcinogenesis (review). Mol. Med. Rep., 6: 923–930.

Guruprasad K., Reddy B.B., Pandit M.W. (1990). Correlation between stability of a protein and its dipeptide composition: a novel approach for predicting *in vivo* stability of a protein from its primary sequence. Protein Eng. Des. Sel., 4: 155–161.

Huang X., Madan A. (1999). CAP3: A DNA sequence assembly program. Genome Res., 9: 868–877.

Imes D.L., Wictum E.J., Allard M.W., Sacks B.N. (2012). Identification of single nucleotide polymorphisms within the mtDNA genome of the domestic dog to discriminate individuals with common HVI haplotypes. Forensic Sci. Int.-Gen., 6: 630–639.

Kim S.K., Lee S.E., Jeong H., Ha J.H. (1998). The complete nucleotide sequence of the domestic dog (*Canis familiaris*) mitochondrial genome. Mol. Phylogenet. Evol., 10: 210–220.

Lowe T.M., Chan P.P. (2016). tRNAscan-SE on-line: search and contextual analysis of transfer RNA genes. Nucleic Acids Res., 44: W54–57.

Łopuszyński W., Szczubiał M., Komsta R. (2010). Prognostic status of p53 protein accumulation in canine malignant mammary tumors. B. Vet. I. Pulawy, 54: 105–111.

McFarland R., Elson J.L., Talyor R.W., Howell N., Turnbull D.M. (2004). Assigning pathogenicity to mitochondrial tRNA mutations: when 'definitely maybe' is not good enough. Genome Anal., 20: P591–596.

Nisztuk-Pacek S., Ślaska B., Grzybowska-Szatkowska L., Babicz M. (2019). Paternal leakage of mitochondrial DNA in the raccoon dog (*Nyctereutes procyonoides* Gray 1834). Ann. Anim. Sci., 19: 61–69.

Panwar B., Raghava G.P.S. (2014). Prediction of uridine modifications in tRNA sequences. BMC Bioinformatics, 15: 326.

Slaska B., Grzybowska-Szatkowska L., Surdyka M., Nisztuk S., Rozanska D., Rozanski P., Smiech A., Orzelski M. (2014). Mitochondrial D-loop mutations and polymorphisms are connected with canine malignant cancers. Mitochondr. DNA, 25: 238–243.

Slaska B., Grzybowska-Szatkowska L., Nisztuk S., Surdyka M., Rozanska D. (2015). Mitochondrial DNA polymorphism in genes encoding ND1, COI and CYTB in canine malignant cancers. Mitochondr. DNA, 26: 452–458.

Surdyka M., Slaska B. (2017 a). Defect of the mitochondrial DNA hypervariable region as a risk factor for canine mammary tumor. Vet. Comp. Oncol., 15: 820–828.

Surdyka M., Slaska B. (2017 b). Defect in ND2, COX2, ATP6, and COX3 mitochondrial genes as a risk factor for canine mammary tumor. Vet. Comp. Oncol., 15: 1062–1072.

Szczubiał M., Łopuszyński W. (2011). Prognostic value of regional lymph node status in canine mammary carcinomas. Vet. Comp. Oncol., 9: 296–303.

Ślaska B., Grzybowska-Szatkowska L., Bugno-Poniewierska M., Surdyka M., Śmiech A. (2013). Nuclear and mitochondrial DNA mutation in human and canine tumors. Med. Weter., 69: 195–202.

Ślaska B., Surdyka M., Brodzki A., Nisztuk S., Gurgul A., Bugno-Poniewierska M., Śmiech A., Różańska D., Orzelski M. (2014). Mitochondrial D-loop mutations can be detected in sporadic malignant tumors in dogs. B. Vet. I. Pulawy, 631–637.

Ślaska B., Grzybowska-Szatkowska L., Bugno-Poniewierska M., Gurgul A., Śmiech A., Różańska D., Dudka J. (2016). Relevance of molecular changes in the *ND4* gene in German Shepherd dog tumors. Pol. J. Vet. Sci., 19: 461–469.

Śmiech A., Ślaska B., Bownik A., Grzybowska-Szatkowska L., Dudka J., Łopuszyński W. (2019). Heteroplasmic mutations and polymorphisms in the *Cyb* gene of mitochondrial DNA in canine mast cell tumors. In Vivo, 33: 57–63.

Śmiech A., Ślaska B., Surdyka M., Grzybowska-Szatkowska L., Łopuszyński W., Różańska D. (2016). Identification of additional mitochondrial DNA mutations in canine mast cell tumors. Acta Vet. Scand., 58: 28.

Tang H., Thomas P.D. (2016). PANTHER-PSEP: predicting disease-causing genetic variants using position-specific evolutionary preservation. Bioinformatics, 32: 2230–2232.

Webb K.M., Allard M.W. (2009). Identification of forensically informative SNPs in the domestic dog mitochondrial control region. J. Forensic Sci., 54: 289–303.

Received: 24 I 2019
Accepted: 26 III 2019