Polish Journal of Veterinary Sciences Vol. 22, No. 1 (2019), 133–141

DOI 10.24425/pjvs.2019.127080

Original article

Characterisation of genome-wide structural aberrations in canine mammary tumours using single nucleotide polymorphism (SNP) genotyping assay

M. Surdyka[1], A. Gurgul[2], B. Slaska[1], K. Pawlina[2], T. Szmatola[2], M. Bugno-Poniewierska[3], A. Smiech[4], K. Kasperek[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] Department of Animal Breeding and Genetics, National Research Institute of Animal Production in Balice, Krakowska 1, 32-083 Balice, Poland
[3] University of Agriculture in Krakow, Institute of Veterinary Science, al. Mickiewicza 24/28 30-059 Krakow, Poland
[4] Sub-Department of Pathomorphology and Forensic Veterinary Medicine, Department and Clinic of Internal Medicine, Faculty of Veterinary Medicine, University of Life Sciences in Lublin, Gleboka 30, 20-950 Lublin, Poland

Abstract

Characterisation of copy number variation (CNV) and loss of heterozygosity (LOH) has provided evidence for the relationship of this type of genetic variation with the occurrence of a broad spectrum of diseases, including cancer lesions. The role of CNVs and germinal or somatic LOHs in canine mammary tumours is still unknown. Therefore, the aim of this study was to identify CNVs and LOHs in canine mammary tumours. Forty-eight samples obtained from normal (n=24) and tumour (n=24) tissues of dogs were analysed. In the study, we used CanineHD BeadChip assay (Illumina) and OncoSNP software to identify copy number alternations in genomes of different dog breeds and in different mammary cancer types occurring in this species. The analyses revealed that, in the case of CNV, the amplification-type variants were longer and more frequent than deletions. Based on the analysis of the frequency of different types of aberrations in the individual parts of the genome, regions that are particularly susceptible to structural aberrations were indicated. The fraction of genes identified within these regions was associated with major processes of neoplastic transformation. Association analysis of such traits as tumour grading as well as the size and age of dogs demonstrated that structural aberrations were more frequent in dogs diagnosed with tumour malignancy grade II and III, in dogs with a larger body size, and in large dogs aged 7-8. The promising results of these pioneering investigations prompt continuation thereof to analyse other types of cancer.

Key words: dog, cancer, mammary tumour, structural aberration, CNV, LOH

Correspondence to: B. Slaska e-mail: brygida.slaska@up.lublin.pl

Introduction

The mammary gland tumour is the second most common cancer in dogs and the most frequent type diagnosed in bitches, accounting for 52% of all diagnosed tumours in these animals. The vast majority of cases of mammary tumours are malignant changes. Breed-related predispositions of dogs for development of cancer diseases suggest an important role of the genetic element in the neoplastic transformation observed in this species. Identification of genetic risk factors may be helpful in prevention and treatment of these diseases and can provide an opportunity to test novel therapeutic strategies (Ślaska et al. 2013, Surdyka and Slaska 2017a,b).

In recent years, there has been considerable progress in the development of tools and techniques required for genomic analyses. This has contributed to identification of high-quality canine genome sequences and development of single-nucleotide polymorphism arrays that are useful in large-scale genomic analyses. Thus, modern genomics offers exceptional potential for genetic characterisation of the background of canine diseases, e.g. cancer, thereby ensuring rapid diagnosis and effective treatment.

A characteristic feature of tumour cells is the large variability of their genomic DNA. Alterations associated with tumours include not only point mutations, small deletions, or insertions but also large-scale structural and ploidy chromosomal aberrations. Genomic aberrations appear in tumours during their evolution promoting a tumour genome that ensures better tumour growth and survival than that of other cells. Large-scale genomic analyses have become a widely used tool for identification of genomic aberrations in human cancer cells, including breast cancer cells (Yang et al. 2013). A similar approach is recently being employed in animal oncogenomics (Gurgul et al. 2014, Pawlina et al. 2016).

Copy number variation (CNV) and loss of heterozygosity (LOH) are two types of genetic structural alterations associated with tumours. CNV is defined as a segment of DNA >= 1kb that is copy-number variable when compared with a reference genome (Feuk et al. 2006). LOH is a common genetic event in cancer development, called in this way because of the early observations of a change in polymorphic markers from a heterozygous state in the germline to an apparently homozygous state in the tumour DNA (Gallie et al. 1983). Genetic alterations include copy number variation (CNV – copy number gains or losses) and loss of heterozygosity (LOH – allelic imbalance or copy-neutral LOH). They may result in inactivation of tumour suppressor genes or activation of oncogenes, which in turn may contribute to uncontrolled growth and metastasis of tumour cells (Perez-Ordonez et al. 2006). Therefore, identification of CNVs and LOHs in tumour samples is indispensable for elucidation of the genetic background, initiation, progression, and differentiation of cancer (Notta et al. 2011). Hence, the aim of the study was to identify and characterise copy number variations and loss of heterozygosity in canine mammary tumours, and to identify genes located in genomic regions characterised by the highest rate of aberrations in the tumours analysed.

Materials and Methods

Samples

The research material consisted of 24 tumour tissues as well as healthy tissue samples collected from the same dogs of various breeds with diagnosed mammary tumours. Prior to DNA extraction, sections for histopathological analysis were sampled from each surgically removed tumour in order to determine the histological type of the tissues analysed. The tissues sampled for histopathological examination were fixed in buffered formalin and embedded in paraffin blocks, and the sections were routinely haematoxylin and eosin (H&E) and toluidine blue. The degree of mammary tumours was assessed using the 3-grade scale (G1- low; G2 – intermediate, G3 – high) (Goldschmidt et al. 2011). DNA was extracted with a commercial DNeasy Blood and Tissue Kit (Qiagen, Hilden, Germany). The isolated genetic material was assessed qualitatively by electrophoretic separation in agarose gel and quantitatively with a BioPhotometer spectrophotometer (Eppendorf, Hamburg, Germany).

Identification of aberrations

The DNA samples were genotyped with the use of a CanineHD Whole-Genome Genotyping assay (Illumina) following the standard Infinium Ultra protocol. The assay contained probes for more than 170,000 single nucleotide polymorphisms with known positions in the CanFam2.0 reference sequence, including 1,547 non-polymorphic intensity-only probes dedicated for CNV (copy number variant) analysis. The SNPs maintained the average density of approx. 70 SNPs per Mb. The data were assessed for quality parameters such as the call rate and log R ratio standard deviation. The minimal requirements for call rates were set at 0.975 and for log R ratio deviation at 0.35 in order to identify samples with poor quality for CNV detection. The aberrations were subsequently detected with the use of the cancer-data dedicated software – OncoSNP

(Yau 2013). It allows detection of copy number variation and copy neutral LOH in tumour samples on the basis of SNP microarray data and identifies only aberrations exclusive for tumour samples, treating intensity data from the healthy tissue as background. Aberrations were called on all autosomes for each control--tumour sample pair according to the manual, applying the GC wave correction as well as -stromal, -intratumour and -allprobes commands (OncoSNP accessed 15.04.2017). All segmental data manipulations (checking for overlapping segments, segments merging and statistics) were performed with PLINK (v1.07) software (Purcell et al. 2007).

Chromosomal fragile sites were identified by choosing regions with a length in the range of 1-10% chromosome coverage and with the number of breakpoints observed within them exceeding 10. The frequency of aberrations for the individual groups (grading, size, age) was calculated as mean aberration frequency in the selected regions of the genome.

Association analysis

The association analyses were carried out in a case--control system for variables: grading, size, and age of dogs classified as large. A simple permutation-based test of association of segmental CNV data for case/control phenotypes was performed in PLINK (1.07) software with 50,000 null permutations generating empirical *p*-values both at the pointwise and genome-wide corrected level. The results were visualised by the "Manhattan plot" using the R package (R version 3.3.2) – a plot of the – log10 (*P*-value) of the association statistic on the y-axis versus the chromosomal position of the SNP on the x-axis. Regions with many highly associated SNPs appeared above the threshold line in the plot.

Identification and functional analysis of genes located in the identified CNVs and LOHs

The identified genomic regions with a variable copy number and known positions in the genome assembly were screened for protein coding genes with the UCSC Genome Browser. Both RefSeq (NCBI) and Ensembl genes were exported and analysed in terms of molecular function of the proteins and biological processes in protein analysis through the evolutionary relationships (PANTHER) classification system (Thomas et al. 2003; PANTHER-Protein Classification System accessed 10.03.2017).

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

Results

Characterisation of identified CNVs

In the biological material with canine mammary tumours, 1823 CNVs generated *de novo* in the tumour tissue were identified. This indicates that approximately 76 CNVs were detected per animal. In case of 58 aberrations, two values for amplification and deletion of DNA fragment were found. Taking into account such CNVs (amplifications or deletions), 1765 unique CNVs were detected. The frequency of individual CNVs was in the range from 4% to 33% in the analysed tumours. CNVs with the highest frequency were localised in position chr34:3023198-45095658. The vast majority of CNVs, i.e. 1725 (97.7%) occurred only once in the examined animals and the other 40 (2.3%) CNVs recurred from 2 to 8 times. The mean CNV length was 3.5 Mb with a median value of 810.8 kb. The analysis revealed the presence of 1488 (84.3%) amplifications, 262 (14.8%) deletions, and 15 (0.9%) alterations combining amplifications and deletions. The mean CNV value of the amplifications was 2-fold higher than that of deletions, i.e. 3 Mb and 1.5 Mb, respectively. A similar relationship was observed for the median values, which were 905.1 kb for the amplifications and 430.9 kb for the deletions.

In the present study, sixteen chromosomal fragile sites were identified. In majority of chromosomal fiagile sites were amplification events, they were located on 11 chromosomes. The mean density of breakpoints located within the regions ranged from 54,409.09 to 98,793.81 bp and was substantially higher than that mean density observed on the individual chromosomes. The fragile sites were identified on chromosomes 6, 8-9, 12-14, 16, 20, 22, and 33-34.

The merger of overlapping unique CNVs into non redundant variable genomic regions (CNVRs) allowed identification of 41 events, covering in total 2,121.88 Mb corresponding to 91.53% of the autosomal genome sequence (2,318.22 Mb) and 86.78% of the whole dog nuclear genome (2,445.11 Mb). The characteristics of CNVRs are shown in Table 1. Among the total CNVRs, there were 14 amplifications and 1 deletion, and the other 26 were copy losses or gains in the same region. The mean sizes of the del/dup events were similar (51.12 Mb and 56.6 Mb, respectively) and substantially greater that the sizes of deletions (0.33 Mb). The median sizes for the del/dup and dup events were similar as well, i.e. 53.05 Mb and 53.65 Mb, respectively.

Germinal LOHs

The analyses revealed the presence of 2,799 germinal LOHs (G-LOHs). Their individual frequency

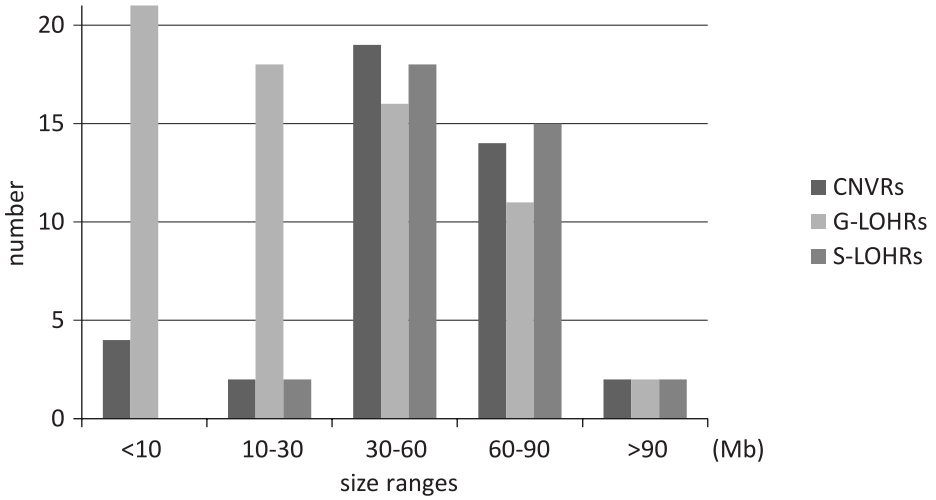

Fig 1. Size distribution of CNVRs, G-LOHRs, and S-LOHRs

ranged from 4.2% to 20.8%. The mean value of hereditary LOHs was 3.25 Mb, whereas the median was 1.35 Mb. The vast majority, i.e. 2,744 (98%) LOHs, were detected only once in the analysed population, and the other 55 (2%) G-LOHs occurred from 2 to 5 times. G-LOH with the highest frequency of 20.8% was identified in position chr3:3593045-4695416. Furthermore, 36 regions with chromosomal fragile sites with higher breakpoint density were identified. The regions were found on 19 chromosomes (usually two regions per chromosome), and the density of the identified breakpoints ranged from 57,185.62 to 99,392.57.

All the identified germinal LOHs were found in 68 unique genomic regions (G-LOHRs). The total length of the identified G-LOHRs was 2096.76 Mb, which accounted for 90.45% of the autosomal genome and 85.75% of the complete nuclear genome of the dog. The mean length of G-LOHRs was 30.83 Mb, and their median value was 24.28 Mb (Table 1).

Somatic LOHs

In the analysed tumour samples, 586 somatic LOHs (S-LOHs) with individual frequency ranging from 4.2% to 33.3% were identified. S-LOHs with the highest frequency were detected on chromosomes 11, 15, 16, 18, and 27. The mean S-LOH size was 5.44 Mb and the median value for the analysed somatic LOHs was 643.2 kb. As many as 91.8% (n=538) of the S-LOH alterations occurred only once in the analysed population. Merely 8.2% (n=48) were somatic LOHs occurring from 2 to 8 times. Chromosomal fragile sites were identified in this analysis variant as well. Three chromosomal fragile sites were situated on chromosomes 9, 14, and 16, and the density of breakpoints was similar to that observed in the two previous analysis variants.

In case of somatic LOHs, there were 37 unique genomic regions exhibiting loss of heterozygosity in tumour cells. The total length of S-LOHRs was 2,154.20 Mb, and the alterations covered 92.92% of the autosomal canine genome and 85.1% of the complete nuclear genome. The mean and median values for S-LOHRs were similar, i.e. 58.22 Mb and 58.24 Mb, respectively (Table 1).

The numbers of CNVRs as well as germinal and somatic LOHRs are compared in Fig. 1. In case of short regions (<10 Mb), the greatest number of alterations were found for the G-LOHR variant, while no such regions were present in the S-LOHR variant. Substantial differences between the number of regions in G-LOHRs and CNVRs and S-LOHRs were found in the 10 do 30 Mb-long regions. In turn, in the range of 30-60 Mb, the number of regions that were susceptible to the alterations was similar in each variant. The presence of two regions with a length >90 Mb was detected in each analysed variant. The longest identified region (122.57 Mb) was the same for CNVRs, G-LOHRs, and S-LOHRs.

Analysis of genes present within CNVRs, G-LOHRs, and S-LOHRs

As revealed by the analysis of genes within CNVRs, G-LOHRs, and S-LOHRs, the greatest number of ensembl genes were identified in case of somatic LOHRs (n=28,225). The ensembl gene content in the other two variants was comparable, i.e. 27,685 for CNVR and 27,504 for G-LOHRs (Table 1). There was no correlation between the number of genes identified on the chromosome and the total length of CNVRs, G-LOHRs, and S-LOHRs localised on the chromosome in any of the analysed cases. Special attention was focused on genes present within the regions of the chromosomal fragile sites in all

Table 1. Summary of CNVRs, G-LOHRs, and S-LOHRs found in this study.

Type	Count	Size (Mb; sum)	Average size (Mb)	Size SD (Mb)	Mediana size (Mb)	Numer of ensembl genes	% reference autosomal genome
CNVR	41	2121.88	51.75	25.61	52.35	27685	91,53
G-LOHR	68	2096.76	30.83	28.04	24.28	27504	90.45
S-LOHR	37	2154.20	58.22	21.24	58.24	28225	92.92

Table 2. Association of CNV, G-LOH, and S-LOH with grading, size, and age.

Variable	Frequency [%]							
	G1	Grading		Size			Age (years)	
		G2	G3	Large	Medium	Small	7-8	12-13
	Variant							
CNV	4.91	13.9	5.63	10.61	2.18	7.80	16.28	6.24
G-LOH	1.98	14.4	7.16	8.81	3.36	9.08	15.35	5.57
S-LOH	4.22	9.24	6.07	12.30	2.52	5.56	13.60	13.18

the analysed variants. In total, 17 reference genes and microRNAs were found in 16 fragile sites identified in the CNV experimental variant; 37 genes were determined in the G-LOH variant and 3 genes were identified for the fragile sites in the S-LOH variant.

Association analysis

Association analysis of CNV, G-LOH, and S-LOH is presented for three variables in the case-control system: grading, size, and age of large-breed dogs (Table 2).

The analysis of the relationship between CNVs and the tumour grade revealed that the genomic sites with higher frequency of CNVs ($p \leq 0.05$) were located on chromosomes 12, 15, 16, and 33. Significant CNVs were observed in 33.3% of G2 individuals in positions chr16:3220145 and chr16:3281580, and in 28.57% in G2 individuals in positions chr16:3281581 and chr16: 3289010. In all these four positions, CNVs were detected in nearly 10% of individuals with G3 tumours. It can be assumed that the higher frequency of this type of aberration in individuals with grade II or III malignancy tumours occurs regularly in canine mammary tumours. The analysis of the association of the germinal LOH frequency with the tumour malignancy grade demonstrated that statistically significant G-LOHs ($p \leq 0.05$) were distributed on as many as 32 chromosomes. Higher frequency of this type of aberration was found in dogs with G2 or G3 tumours. Alterations on chromosomes 25, 31, and 34 were detected in 28.57% of dogs with G2 tumours and alterations in position chr34: 4958390in were noted in 14.29% of individuals with G3 tumours. In this case, as in the two other aberration variants, the highest frequency of significant S-LOHs ($p \leq 0.05$) was reported in dogs with G2 tumours. In several positions on chromosome 16, significant aberrations ($p \leq 0.05$) were detected in 42.86% of dogs with G2 tumours. The frequency of statistically significant S-LOHs in G3 tumour samples was 14.29%, whereas no significance was found for the lowest malignancy grade.

The analysis performed in the group of dogs divided according to the body size revealed the presence of 34 CNVs with statistically significant differences ($p \leq 0.05$) in the frequency in each group. The frequency of these aberrations was shown to be higher in large dogs, compared with medium-sized dogs. No significant differences in the frequency of the aberrations were found in the group of small dogs with the exception of four positions on chromosome 12, where CNVs were detected in 33.3% of the dogs. In case of G-LOHs, 353 significant aberrations were identified with the highest frequency among the large dogs. In positions chr4:7310265, chr4:7935673, chr4:7966609, chr15:3657068, and ch15:4031228, G-LOHs were identified in 37.5% of large dogs, and in positions chr4:7935673 and chr4:7966609 alterations were detected in 12.5% of medium-sized dogs. In case of somatic LOHs, only a few statistically significant differences ($p \leq 0.05$) in the frequency of these aberrations were revealed in the groups of dogs distinguished according to their body size. S-LOHs were observed more frequently in the large or medium-sized dogs than in the small dogs. The highest frequency was found in positions

chr5:3037864 and p5:3110350, i.e. 8.33% for the large dogs and 20.83% for the medium-sized dogs.

The analysis performed in the two age groups resulted in identification of 40 genomic loci exhibiting statistically significant differences ($p \leq 0.05$) in the frequency of CNVs in the two age groups of large dogs. The frequency of these aberrations was higher in the 7-8 year-olds. Similarly, in case of germinal LOHs, the frequency of significant aberrations was higher in the younger dogs (7-8 years old). In only two positions on chromosome 3 and three positions on chromosome 16, significant G-LOHs were identified in 55.56% and 44.44% of 12-13-year-old dogs, respectively. In case of the S-LOH variant, the frequency of this type of aberrations in the two age groups did not differ significantly between the groups. This correlation is not surprising, since the average lifespan of large breeds is relatively short and often does not exceed 10 years. Therefore, the age of 7-8 years in large dogs can be referred to as old or senile. This may explain the higher frequency of the statistically significant aberrations ($p \leq 0.05$) in the large dogs aged 7-8 than in the 12-13-year olds.

Discussion

Copy number variation, which is one of sources of the genetic diversity, has become the target of intensive research in modern genomics. Genomic variations in tumour cells are complex and their variants may be related to susceptibility to different types of cancer (Notta et al. 2011). Investigation of genomic aberrations in tumours is therefore an important step towards identification of genetic abnormalities as somatic patterns associated with the initiation and progression of tumours, including canine mammary cancer. In the analysed biological material 1765 unique CNVs (amplifications or deletions) were identified. 84.3% unique CNVs were amplifications, 14.8% deletions and and 15 (0.9%) alterations combining amplifications and deletions. Dominance of amplification over deletion events, in terms of both the quantity and length, was observed in human tumours, including mammary and breast cancers (O'Hagan et al. 2002). There are reports of high frequency of amplification events in certain chromosomal regions in female breast cancer patients (Orsetti 1999). These regions are often referred to as hotspots of neoplastic transformation. The origin and conservation of such regions in the genome is still unclear. Presumably, the first genome rearrangements resulting in amplification take place already at the onset of the neoplastic process; hence, numerous and long-term amplification alterations may accumulate over a relatively long period, leading to conservation thereof in the genome of tumour cells. Consequently, their important role as sites of neoplastic transformation should be acknowledged.

The considerable dominance of amplifications over deletions and their greater length observed in the present study may indicate a greater impact of this type of alterations on the genome. Amplifications are more ambiguous than deletions because the extra copies of DNA may end up elsewhere in the genome and may be affected by subsequent events in those regions. It is also highly probable that the predominance of amplification over deletion events is an effect of the transformation of a healthy cell into a tumour cell. Gene amplification has been shown to be a typical genetic disorder in case of cancer, and many oncogenes have been identified in genomic regions that are amplification sites. Current data indicate that tumours often develop as a result of disorders in several oncogenes. It has been demonstrated that co-amplified genes with known oncogenes present within the same amplicon promote tumour development. Therefore, identification of genes in genomic regions where an amplification has taken place can complement the current knowledge. The dominance of amplification over other genomic aberrations has been investigated for over 20 years. However, why some genomic regions are more susceptible to amplification than the others has not been fully elucidated yet (Wahl et al. 1983). The propensity of certain genomic regions to amplify is still a frequent subject of researchers' interest. Genomic sequences that deserve special attention are the chromosomal fragile sites, which tend to form constrictions and breakpoints when the cell is exposed to replication stress. Over 100 fragile sites have been identified in the human genome to date. Fragile sites exhibit high heterogeneity and their length in the human genome ranges from 0.3 to 9 Mb. Such regions have been identified in various types of cancer (Glover et al. 2005). They are candidate regions that can initiate amplification through their proneness to breakage. In the present study, sixteen chromosomal fragile sites were identified and they were located on 11 chromosomes. The chromosomal fragile sites were in the majority amplifications events.

In investigations conducted by (Gurgul et al. 2014) in a group of 26 dogs of different breeds with different types of cancer, 426 unique chromosomal aberrations were identified in tumour tissues, including 273 deletions and 153 amplifications to 4 copies. Other studies of canine genomics revealed that deletions were more numerous than duplications in this type of analysis. It was found that although deletions are smaller than duplications, their higher frequency can affect the

genome equally as duplications (Chen et al. 2009, Conrad et al. 2010, Berglund et al. 2012).

In case of cancer, loss of heterozygosity is more frequent than copy number variation, as it causes irreversible loss of one of the parental alleles. In turn, copy gains are often unstable. Microarray analyses are applied in oncogenomics studies for determination of the distribution of the LOH length across the genome (Abkevich et al. 2012). Loss of heterozygosity in tumour cells most often results from deletion of a certain chromosomal region or an entire chromosome; it is frequently accompanied by reduction of the copy number. Abkevich et al. (2012) hypothesised that various length LOH regions may appear in the tumour cell genome via different pathways. Correlations were determined between the number of short LOHs (<15Mb), the number of long LOH regions (>15Mb, but smaller than the entire chromosome), and the number of LOH regions covering entire chromosomes and homologous recombination (HR) deficiency. Special attention was devoted to somatic and congenital defects in genes *BRCA1* and *BRCA2* supporting homologous recombination, which is associated with predisposition to various cancers, e.g. mammary and ovarian cancers (Venkitaraman 2002). The investigation results indicate that the medium size class of LOH regions >15 MB but smaller than the entire chromosome is positively correlated with defects in the HR pathway genes. No such correlation was found between very short LOH regions and those covering the entire chromosome.

Analysis of genes present within CNVRs, G-LOHs and S-LOHs, revealed numerous of ensemble genes were identified in all the analysed variants (Table 1). Special attention was focused on genes present within the regions of the chromosomal fragile sites in all the analysed variants. As reported in the literature, structural abnormalities in four of all genes, i.e. *ZFPM2*, *EBAG9*, *AQP1*, and *MSTN*, have an impact on disorders of biological functions commonly associated with neoplastic transformation. Protein products of these genes are associated with disturbances in apoptotic cell death in dogs (UniProt accssed 10.03.2017). Negative regulation of the apoptotic process in dogs was detected in case of gene *AQP1* and, as evidenced repeatedly, any abnormalities in apoptosis play a key role in neoplastic transformation (Pawłowski et al. 2013). Furthermore, it has been shown that the protein product of gene *AQP1* is involved in the cellular response to stress or maintenance of homeostasis, i.e. in biological cell functions that are disturbed during cancer development (UniProt accessed 10.03.2017). Therefore, the aberrations in genes *ZFPM2*, *EBAG9*, *AQP1*, and *MSTN* may have resulted in disturbances in the function of proteins encoded by the genes and disorders in such processes as apoptosis, homeostasis, or stress response, thereby indirectly contributing to the development of mammary tumours in the analysed dogs.

The selected limited group of CNVs and germinal and somatic LOHs characterised by the highest frequency in the examined population and potentially greater importance in the group of dogs with mammary tumour facilitated identification of several microRNA molecules, which play a significant role in the development of many human tumours, as indicated by recent literature data. The role of these molecules in cancer diseases in various animal species, including dogs, has been poorly explored to date. Three canine microRNA molecules (MIR875, MiR448, MIR-15a), whose human counterparts are involved in the neoplastic transformation process in the case of gastric, brain, or lung cancer, were determined in the examined group (Braoudaki et al. 2016).

Association analysis of CNV, G-LOH, and S-LOH is presented for three variables in the case-control system: grading, size, and age of large-breed dogs (Table 2). Due to the high incidence of mammary tumours in dogs, the risk factors associated with the higher prevalence of this type of cancer have been investigated by researchers for over 40 years. An increased incidence of canine mammary tumour has been found after the 6th year of dog's life and maximally between year 9 and 11. Mammary tumours are rarely detected in dogs younger than 2 years of age (Schneider 1970). A higher risk of this type of cancer has also been demonstrated in large rather than small dogs, and the Boxer, mongrel, and Chihuahua, i.e. small or medium-sized breeds, have been classified as the least prone breeds (Brodey et al. 1983, Zatloukal et al. 2005). Association analyses are also carried out in terms of the tumour malignancy grade. The histological classification of canine mammary tumours distinguishes three malignancy degrees: grade I, II, and III. Karayannopoulou et al. (2005) showed differences in the survival of dogs with various malignancy grades. The survival rate in dogs with grade III malignant neoplasia was lower than in the case of grades I and II. The higher grade was also associated with worse prognosis and higher metastasizing capability (Chang et al. 2007). These observations have been confirmed by analysis carried out with the latest molecular biology techniques and detailed statistical tests. Another analysis (Surdyka and Slaska 2017a,b) conducted on the group of dogs examined in this study demonstrated a statistically significant correlation between changes in the mitochondrial genome and larger body size and older age. However, no association with grading was then found.

The results of the analyses indicate that higher frequency of structural aberration in all analysed variants were observed in individuals with grade II and III malignancy tumours. The regularity observed in the association of structural aberrations with the tumour malignancy grade confirms the histopathological observations reported to date (Chang et al. 2007). It is therefore probable that some structural aberrations appear only in tumours with a higher malignancy grade or they are the cause of an increase in the grade of canine mammary tumours.

The analysis performed in the group of dogs divided according to the body size revealed that frequency of CNVs, G-LOHs and S-LOHs was higher in large dogs compared with medium or small-sized dogs (Table 2). The correlations presented in this study are consistent with long-term observations reported by many research teams (Brodey et al. 1983, Zatloukal et al. 2005).

The analysis performed in the two age groups showed that in case of CNVs and G-LOHs, the frequency of structural aberrations was higher in the 7-8 years-olds dogs. In case of the somatic LOHs, the frequency of this type of aberrations was similar in the two age groups. These observations are consistent with those reported by Schneider (1970) and Brodey et al. (1983).

Conclusion

The application of the SNP genotyping assay and the OncoSNP calculation tool yielded detailed characterisation of structural aberrations that are exclusively specific to canine mammary tumour cells. The results of the analyses indicate that the amplifications were longer and more frequent than the deletions in all the experimental variants. Genomic regions with high frequency of aberrations were identified as well. These were 19 regions identified for the CNV variants, 45 regions for the germinal LOHs, and 4 regions for the somatic LOHs. Additionally, genes in regions that were especially susceptible to these rearrangements were identified. Some of them were associated with processes that are important for neoplastic transformation. The association analysis performed for the grading, size, and age variants allows theconclusion that structural aberrations are more frequent in dogs with grade II and III malignancy, dogs with a larger body size, and large dogs aged 7-8 years.

Our research is the first attempt at SNP microarray-based analysis of CNV and LOH in the genomes of dogs with mammary tumours. In the study, we have focused on identification of genomic regions that are particularly susceptible to severe structural aberrations, which are likely to be associated with neoplastic transformation. The promising results of these pioneering investigations prompt continuation of the research and analysis of other types of cancer.

References

Abkevich V, Timms KM, Hennessy BT, Potter J, Carey MS, Meyer LA, Smith-McCune K, Broaddus R, Lu KH, Chen J, Tran TV, Williams D, Iliev D, Jammulapati S, FitzGerald LM, Krivak T, DeLoia JA, Gutin A, Mills GB, Lanchbury JS **(2012)** Patterns of genomic loss of heterozygosity predict homologous recombination repair defects in epithelial ovarian cancer. Br J Cancer 107: 1776-1782.

Berglund J, Nevalainen EM, Molin AM, Perloski M; LUPA Consortium, André C, Zody MC, Sharpe T, Hitte C, Lindblad-Toh K, Lohi H, Webster MT **(2012)** Novel origins of copy number variation in the dog genome. Genome Biol 13: R73.

Braoudaki M, Lambrou GI, Giannikou K, Papadodima SA, Lykoudi A, Stefanaki K, Sfakianos G, Kolialexi A, Tzortzatou-Stathopoulou F, Tzetis M, Kitsiou-Tzeli S, Kanavakis E **(2016)** miR-15a and miR-24-1 as putative prognostic microRNA signatures for pediatric pilocytic astrocytomas and ependymomas. Tumour Biol 37: 9887-9897.

Brodey RS, Goldschmidt MH, Roszel JR **(1983)** Canine mammary gland neoplasms. J Am Anim Hosp Assoc 19: 61-90.

Chang SC, Chang CC, Chang TJ, Wong ML **(2007)** Prognostic factors associated with survival two years after surgery in dogs with malignant mammary tumours: 79 cases (1998-2002) J Am Vet Med 227: 1625-1629.

Chen WK, Swartz JD, Rush LJ, Alvarez CE **(2009)** Mapping DNA structural variation in dogs. Genome Res 19: 500-509.

Conrad DF, Pinto D, Redon R, Feuk L, Gokcumen O, Zhang Y, Aerts J, Andrews TD, Barnes C, Campbell P, Fitzgerald T, Hu M, Ihm CH, Kristiansson K, Macarthur DG, Macdonald JR, Onyiah I, Pang AW, Robson S, Stirrups K, Valsesia A, Walter K, Wei J, Wellcome Trust Case Control Consortium, Tyler-Smith C, Carter NP, Lee C, Scherer SW, Hurles ME **(2010)** Origins and functional impact of copy number variation in the human genome. Nature 464: 704-712.

Gallie BL, Murphree AL, Strong LC, White RL **(1983)** Expression of recessive alleles by chromosomal mechanisms in retinoblastoma. Nature 305: 3134.

Glover TW, Arlt MF, Casper AM, Durkin SG **(2005)** Mechanisms of common fragile site instability. Hum Mol Genet 14 (Suppl 2): R197-R205.

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

Gurgul A, Żukowski K, Ślaska B, Semik E, Pawlina K, Ząbek T, Bugno-Poniewierska M **(2014)** General assessment of copy number variation in normal and tumor tissues of the domestic dog (Canis lupus familiaris). J Appl Genet 55: 353-363.

Karayannopoulou M, Kaldrymidou E, Constantinidis TC, Dessiris A **(2005)** Histological grading and prognosis in

dogs with mammary carcinomas: application of a human grading method. J Comp Pathol 133: 246-252.

Notta F, Mullighan CG, Wang JC, Poeppl A, Doulatov S, Dick JE **(2011)** Evolution of human BCR-ABL1 lymphoblastic leukaemia-initiating cells. Nature 469: 362-367.

O'Hagan RC, Chang S, Maser RS, Mohan R, Artandi SE, Chin L, DePinho RA **(2002)** Telomere dysfunction provokes regional amplification and deletion in cancer genomes. Cancer Cell 2.2: 149-155.

OncoSNP, **(2017)** https://sites.google.com/site/oncosnp/home [accessed 15.04.2017]. Orsetti B, Courjal F, Cuny M, Rodriguez C, Theillet C **(1999)** 17q21-q25 aberrations in breast cancer: combined allelotyping and CGH analysis reveals 5 regions of allelic imbalance among which two correspond to DNA amplification. Oncogene 18: 6262-6270.

PANTHER – Protein Classification System (2017) http://www.pantherdb.org [accessed 10.03.2017].

Pawlina K, Gurgul A, Bugno-Poniewierska M, Klukowska-Rötzler J, Koch C, Mählmann K **(2016)** P6041 Microarray analysis of genomic aberrations of horse sarcoids. J Anim Sci 94 (Suppl 4):169.

Pawłowski KM, Mucha J, Majchrzak K, Motyl T, Król M **(2013)** Expression and role of PGP, BCRP, MRP1 and MRP3 in multidrug resistance of canine mammary cancer cells. BMC Vet Res 9: 119.

Perez-Ordonez B, Beauchemin M, Jordan RC **(2006)** Molecular biology of squamous cell carcinoma of the head and neck. J Clin Pathol 59: 445-453.

Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Sham PC **(2007)** PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 81: 559-575.

Schneider R **(1970)** Comparison of age, sex, and incidence rates in human and canine breast cancer. Cancer 26: 419-426.

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

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

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.

Thomas PD, Campbell MJ, Kejariwal A, Mi H, Karlak B, Daverman R, Diemer K, Muruganujan A, Narechania A **(2003)** PANTHER: a library of protein families and subfamilies indexed by function. Genome Res 13: 2129-2141. UniProt (2017) http://www.uniprot.org [accessed 10.03.2017].

Venkitaraman AR **(2002)** Cancer susceptibility and the functions of BRCA1 and BRCA2. Cell 208: 171-182.

Wahl GM, de Saint Vincnet BR, DeRose ML **(1983)** Effect of chromosomal position on amplification of transfected genes in animal cells. Nature 307: 516-520.

Yang L, Liu B, Huang B, Deng J, Li H, Yu B, Qiu F, Cheng M, Wang H, Yang R, Yang X, Zhou Y, Lu J **(2013)** A functional copy number variation in the WWOX gene is associated with lung cancer risk in Chinese. Hum Mol Genet 22: 1886-1894.

Yau C **(2013)** OncoSNP-SEQ: a statistical approach for the identification of somatic copy number alterations from next-generation sequencing of cancer genomes. Bioinformatics, 29: 2482-2484.

Yau C, Mouradov D, Jorissen RN, Colella S, Mirza G, Holmes CC **(2010)** A statistical approach for detecting genomic aberrations in heterogeneous tumor samples from single nucleotide polymorphism genotyping data. Genome Biol 11: R92.

Zatloukal J, Lorenzova J, Tichý F, Nečas A, Kecova H, Kohout P **(2005)** Breed and age as risk factors for canine mammary tumours. Acta Vet Brno 74: 103-109.