Molecular Homology and Difference between Spontaneous Canine Mammary Cancer and Human Breast Cancer

Deli Liu1, Huan Xiong1, Angela E. Ellis2, Nicole C. Northrup2, Carlos O. Rodriguez Jr3, Ruth M. O’Regan4, Stephen Dalton1, and Shaying Zhao1

Abstract

Spontaneously occurring canine mammary cancer represents an excellent model of human breast cancer, but is greatly understudied. To better use this valuable resource, we performed whole-genome sequencing, whole-exome sequencing, RNA-seq, and/or high-density arrays on twelve canine mammary cancer cases, including seven simple carcinomas and four complex carcinomas. Canine simple carcinomas, which histologically match human breast carcinomas, harbor extensive genomic aberrations, many of which faithfully recapitulate key features of human breast cancer. Canine complex carcinomas, which are characterized by proliferation of both luminal and myoepithelial cells and are rare in human breast cancer, seem to lack genomic abnormalities. Instead, these tumors have about 35 chromatin-modification genes down-regulated and are abnormally enriched with active histone modification H4-acetylation, whereas aberrantly depleted with repressive histone modification H3K9me3. Our findings indicate the likelihood that canine simple carcinomas arise from genomic aberrations, whereas complex carcinomas originate from epigenomic alterations, reinforcing their unique value. Canine complex carcinomas offer an ideal system to study myoepithelial cells, the second major cell lineage of the mammary gland. Canine simple carcinomas, which faithfully represent human breast carcinomas at the molecular level, provide indispensable models for basic and translational breast cancer research. Cancer Res; 74(18): 5045–56. ©2014 AACR.

Introduction

Spontaneous cancers in pet dogs represent one of the best cancer models (1–8). First, these cancers are naturally occurring and heterogeneous, capturing the essence of human cancer, unlike genetically modified or xenograft rodent models. Second, as companion animals, dogs share the same environment as humans and are exposed to many of the same carcinogens. Indeed, environmental toxins, advancing age, and obesity are also risk factors for canine cancer (1). Third, dogs better resemble humans in biology, for example, similar telomere and telomerase activities (9) and frequent spontaneous carcinogens. Indeed, environmental toxins, advancing age, and obesity are also risk factors for canine cancer (1). Third, dogs better resemble humans in biology, for example, similar telomere and telomerase activities (9) and frequent spontaneous alterations, reinforcing their unique value. Canine complex carcinomas offer an ideal system to study myoepithelial cells, the second major cell lineage of the mammary gland. Canine simple carcinomas, which faithfully represent human breast carcinomas at the molecular level, provide indispensable models for basic and translational breast cancer research. Cancer Res; 74(18): 5045–56. ©2014 AACR.

Like other cancer types, many anatomic and clinical similarities are documented between canine mammary cancer and human breast cancer (24). Various molecular homologies are also reported, for example, WNT signaling alteration (15, 17). Meanwhile, canine mammary cancer also differs from human breast cancer in certain aspects. For example, dogs have only population of pet dogs (~70 million estimated in the United States) provides a valuable resource facilitating basic and clinical research. Importantly, the dog genome has been sequenced to a >7.6X coverage (11), yielding a genome assembly nearly as accurate as the mouse or rat genome (11, 12), unlike another companion animal, the cat. This makes many genomic analyses possible with the dog but not with the cat.

As in women, mammary cancer is among the most frequent cancers in female dogs. The annual incidence rate is estimated at 198 per 100,000 (1), which is comparable with the rate of 125 per 100,000 for breast cancer in women in the United States (13). Mammary cancer is especially common in dogs that are not spayed or are spayed after the second estrus, with the risk for malignant tumor development expected at 26% (1). However, unlike human breast cancer, canine mammary cancer is poorly characterized at the genome-wide level. For example, only five canine mammary cancer cases have recently undergone approximately 2X whole-genome sequencing (WGS; ref. 14), and a limited number have been analyzed with gene expression microarray (15–17). This drastically differs from their human counterparts, where thousands of breast cancer genomes and transcriptomes are characterized, with several studies cited here (18–23).

Like other cancer types, many anatomic and clinical similarities are documented between canine mammary cancer and human breast cancer (24). Various molecular homologies are also reported, for example, WNT signaling alteration (15, 17). Meanwhile, canine mammary cancer also differs from human breast cancer in certain aspects. For example, dogs have only population of pet dogs (~70 million estimated in the United States) provides a valuable resource facilitating basic and clinical research. Importantly, the dog genome has been sequenced to a >7.6X coverage (11), yielding a genome assembly nearly as accurate as the mouse or rat genome (11, 12), unlike another companion animal, the cat. This makes many genomic analyses possible with the dog but not with the cat.

As in women, mammary cancer is among the most frequent cancers in female dogs. The annual incidence rate is estimated at 198 per 100,000 (1), which is comparable with the rate of 125 per 100,000 for breast cancer in women in the United States (13). Mammary cancer is especially common in dogs that are not spayed or are spayed after the second estrus, with the risk for malignant tumor development expected at 26% (1). However, unlike human breast cancer, canine mammary cancer is poorly characterized at the genome-wide level. For example, only five canine mammary cancer cases have recently undergone approximately 2X whole-genome sequencing (WGS; ref. 14), and a limited number have been analyzed with gene expression microarray (15–17). This drastically differs from their human counterparts, where thousands of breast cancer genomes and transcriptomes are characterized, with several studies cited here (18–23).

Like other cancer types, many anatomic and clinical similarities are documented between canine mammary cancer and human breast cancer (24). Various molecular homologies are also reported, for example, WNT signaling alteration (15, 17). Meanwhile, canine mammary cancer also differs from human breast cancer in certain aspects. For example, dogs have only population of pet dogs (~70 million estimated in the United States) provides a valuable resource facilitating basic and clinical research. Importantly, the dog genome has been sequenced to a >7.6X coverage (11), yielding a genome assembly nearly as accurate as the mouse or rat genome (11, 12), unlike another companion animal, the cat. This makes many genomic analyses possible with the dog but not with the cat.

As in women, mammary cancer is among the most frequent cancers in female dogs. The annual incidence rate is estimated at 198 per 100,000 (1), which is comparable with the rate of 125 per 100,000 for breast cancer in women in the United States (13). Mammary cancer is especially common in dogs that are not spayed or are spayed after the second estrus, with the risk for malignant tumor development expected at 26% (1). However, unlike human breast cancer, canine mammary cancer is poorly characterized at the genome-wide level. For example, only five canine mammary cancer cases have recently undergone approximately 2X whole-genome sequencing (WGS; ref. 14), and a limited number have been analyzed with gene expression microarray (15–17). This drastically differs from their human counterparts, where thousands of breast cancer genomes and transcriptomes are characterized, with several studies cited here (18–23).
one or two estrous cycles a year followed by a prolonged luteal phase, which is 63 days for the dog compared with 14 days for the human. During this unusually long luteal phase, the mammary gland is continuously exposed to high levels of progesterone (24). Another variation that may be related to this hormonal difference is described below.

Mammary gland epithelium consists of an inner layer of luminal cells and an outer layer of myoepithelial cells that border the basal lamina. Compared with luminal cells, myoepithelial cells are significantly understudied and poorly characterized (25–30). Although their importance in mammary gland development and pathogenesis has been noted (26, 29, 31), myoepithelial cells have traditionally received far less attention than luminal cells. This is at least partially because they rarely proliferate in human breast cancer (32, 33). However, in canine mammary cancer, myoepithelial cell proliferation is much more common, occurring in >20% canine tumors compared with <0.1% human tumors (34). To more effectively use this unique feature of canine mammary cancer for a better understanding of myoepithelial cells, we set out to comprehensively compare spontaneous canine mammary cancers with and without myoepithelial cell proliferation and to evaluate their molecular similarities to and differences from human breast cancers.

Materials and Methods

Canine mammary tissue samples

Fresh-frozen (FF) and formalin-fixed paraffin-embedded (FFPE) normal and tumor tissue samples of spontaneous canine mammary cancer were obtained from the University of California-Davis School of Veterinary Medicine (Davis, CA) and the Animal Cancer Tissue Repository of the Colorado State University (Fort Collins, CO). Samples were collected from client-owned dogs that develop the disease spontaneously, under the guidelines of the Institutional Animal Care and Use Committee and with owner informed consent. The breed, age, histopathologic descriptions, and other information are provided in Supplementary Table S1.

Immunohistochemical analyses

Immunohistochemical (IHC) experiments were performed following standard protocols with 5-µm FFPE sections. Primary antibodies were used as described (35), including those against smooth muscle myosin heavy chain clone ID8 (MAB3568), acetyl-H4 (06-866), H3K9me3 (07-442), and H3-K4/K9-me3 (06-866), all from Millipore; H3K4me3 (Abcam; ab8580), estrogen receptor alpha clone E115 (Abcam; ab32063); and E-cadherin (R&D Systems; AF648). Alexa Fluor488–, 647– or 594–conjugated secondary antibodies are from Jackson ImmunoResearch. Images were taken with a Zeiss LSM 710 confocal microscope.

Tissue dissection, DNA and RNA extraction, and PCR analyses

Cryosectioning of FF tissues, hematoxylin and eosin (H&E) staining, and cryomicrodissection were performed as described (5) to enrich tumor cells for tumor samples and mammary gland epithelial cells for normal samples. Genomic DNA and RNA were then extracted from the dissected tissues using the DNeasy Blood & Tissue Kit (cat. no. 69504), RNeasy Plus Mini Kit (cat. no. 74134), or AllPrep DNA/RNA Mini Kit (cat. no. 80204) from Qiagen. Only samples with a 260/280 ratio of approximately 1.8 (DNA) or approximately 2.0 (RNA) and showing no degradation and other contaminations on the agarose gels were subjected to further analyses. The synthesis of cDNA, primer design, and PCR or qPCR with genomic DNA or cDNA samples were conducted as described (6). Primers used are listed in the Supplementary Methods.

Array comparative genomic hybridization analyses

Array comparative genomic hybridization (aCGH) experiments were conducted at the Florida State University Microarray Facility, with 385 K canine CGH array chips from Roche NimbleGen Systems, Inc. Copy-number abnormalities (CNA) were identified as described (5).

Paired-end WGS, whole-exome sequencing, and RNA-seq

All three types of sequencing were conducted using the Illumina platform, following the protocols from the manufacturer. Paired-end WGS of >12X sequence coverage was performed in collaboration with the Emory Genome Center (50 bp or 100 bp paired-end sequencing of ~200 bp fragments) or the BGI-America (90 bp paired-end sequencing of ~500 bp fragments). Whole-exome sequencing (WES) was conducted in collaboration with the Hudsonalpha Institute for Biotechnology (Huntsville, AL). First, exome-capturing was achieved by using a solution-based SureSelect Kit from Agilent, covering 50 Mb canine exons and adjacent regions. Then, paired-end sequences of 50 bp of approximately 200-bp fragments were generated from the captured targets to reach the coverage of 134X to 245X. RNA-seq was performed at Hudsonalpha, yielding 42 to 94 million paired-end sequence reads of 50 bp per sample.

Sequence data analyses

Sequence read alignment, mutation discovery, translocation and chimeric fusion gene identification, clustering, and other analyses are provided in the Supplementary Methods. Briefly, WGS, WES, and RNA-seq sequence reads were aligned to the dog reference genome (11). Then, uniquely mapped WES reads were used to detect base substitutions and small indels, and significantly mutated genes were identified as described (20). Uniquely mapped WGS read pairs were used to identify somatic translocations and chimeric fusion genes. Uniquely mapped RNA-seq reads were used to quantify each gene’s expression level, as well as to detect chimeric fusion transcripts and sequence mutations.

Results

Canine simple carcinomas have no myoepithelial cell proliferation, whereas canine complex carcinomas have luminal and myoepithelial cells, both proliferating

The 12 cases subjected to genome-wide characterization represent two major histologic subtypes of canine mammary cancer (34); five with myoepithelial cell proliferation (complex carcinomas) and 7 without (simple carcinomas;
Supplementary Table S1). Tumor cells in simple carcinomas express only luminal markers such as E-cadherin (Fig. 1A), and histologically match typical human breast carcinomas (Fig. 1A and C). Tumors with myoepithelial cell proliferation include four complex carcinomas, a subtype that is rare in humans (32), and one carcinoma with two distinct histologic regions, one considered simple and the other considered complex. Complex carcinomas have prominent expression of both the luminal marker E-cadherin and the myoepithelial marker smooth muscle myosin heavy chain (SMHC; Fig. 1B), indicating dual proliferation of luminal and myoepithelial cells. This is also visible in H&E-stained sections (Fig. 1D). Besides this histologic difference, the tumors also vary in cancer progression stages (in situ, invasive, or metastatic to the lung) and in estrogen receptor (ER) expression (five ER+ tumors and seven ER- tumors; Supplementary Table S1).

**CNAs are frequent in canine simple carcinomas**

Reminiscent of human breast cancer, canine simple carcinoma genomes harbor extensive CNAs. First, we observed both focal and broad CNAs totaling from 10 Mb to >100 Mb and affecting hundreds of genes per tumor (Fig. 2A and B; Supplementary Tables S2A–S2C), with the extensiveness of CNA in correlation with the tumor progression stage. The only exception to this is an inflammatory carcinoma, in which no CNAs were detected. Second, CNAs also occurred in genomic sites recurrently altered in human breast cancer, for example, amplification of human 8q and dog chromosome 13 that encode genes including MYC (Fig. 2A). Third, although large deletions were discovered, one resulting in PTEN loss (Fig. 2A; Supplementary Table S2G), amplifications prevailed over deletions in most tumors. Notably, two large amplicons of >4 Mb, harboring 54 and 43 genes, respectively, were uncovered (Fig. 2C). This led to amplification and overexpression of oncogenes such as BRAF, PIM1, and CCND3 (Supplementary Tables S2E and S2F).

**Translocations and a superamplicon were discovered in a canine simple carcinoma by paired-end WGS**

To further explore the two >4 Mb amplicons described above, we sequenced the tumor and normal genomes of case 76 (Fig. 2A) to a >15X sequence coverage (Supplementary Fig. S1 and Supplementary Table S2D). For comparison purposes, similar sequencing was performed on the case having the most extensive CNAs (case 406434, with pulmonary metastasis) and
another case having hardly any CNAs (case 32510). WGS revealed fewer translocations and inversions than CNAs in these tumors. Furthermore, reminiscent of the human breast cancer MCF7 genome (36), some translocations are associated with amplification, creating a superamplicon with loci from different chromosomes colocalized and coamplified (Fig. 2D).

On the basis of chimeric sequence reads that span the translocation junctions (Supplementary Table S2H) and PCR confirmation (Supplementary Fig. S1), we propose a mechanism for the superamplicon formation (Fig. 2D). First, a circle, consisting of approximately 1-Mb sequences from chromosome 12 and approximately 0.4 Mb from chromosome 16, emerged via two translocations that were likely facilitated by prior sequence amplification. The circle, which harbors oncogene PIM1 and 17 other genes (Supplementary Tables S2E and S2F), was then further amplified.

The superamplicon harbors a potentially oncogenic fusion gene, ZFAND3–MGAM, created via a translocation

The superamplicon also harbors a newly created fusion gene. It consists of the first four exons of ZFAND3, a zinc finger gene located on chromosome 12, and the last 22 to 49 exons of MGAM, which encodes maltase-glucoamylase and is located on chromosome 16 (Fig. 2F). The fusion gene, termed ZFAND3-MGAM, arose from a translocation occurring in introns; transcription and splicing then yielded an in-frame fusion transcript. This was confirmed by the detection of chimeric sequence fusion points via WGS, RNA-seq, and PCR.
analyses (Supplementary Fig. S1 and Supplementary Table S2H).

As a result of in-frame fusion, the A20-type zinc finger domain of ZFAND3 and the glucoamylase domain of MGAM are kept intact in the fusion product (Fig. 2F). This seems to be significant. First, the A20 zinc finger protein has been reported to inhibit tumor necrosis factor–induced apoptosis (37). Second, MGAM, an integral membrane protein with its catalytic domains facing the extracellular environment, is normally expressed in the intestine to digest starch into glucose (38). Indeed, we did not detect MGAM expression in normal mammary tissues, unlike ZFAND3 (Supplementary Fig. S1C). However, in carcinoma 76, both MGAM and ZFAND3–MGAM are amplified and overexpressed. ZFAND3–MGAM, which lacks the transmembrane domain and becomes intracellular, could promote oncosgenesis by accelerating carbohydrate metabolism via its glucoamylase domain and meanwhile inhibiting apoptosis via its A20-type domain. Of course, whether this is true or not requires further studies.

**Somatic sequence mutations are frequent in canine simple carcinomas as revealed by WES**

To examine somatic base substitutions and small indels, we performed WES on the matching tumor and normal genomes of four simple carcinoma cases to 134X to 245X coverage (Supplementary Fig. S2 and Supplementary Table S3A). This analysis again revealed several dog–human homologies. First, base transitions, particularly C → T/G → A changes, dominate base transversions in most tumors (Fig. 3A), indicating similar mutation mechanisms (e.g., deamination of 5mC to T) in both the species. The only exception (tumor 406434) has C → A/G → T transversions predominating, which is not an experimental artifact of WES (39) based on our analyses (Supplementary Table S3F), and concurrently harbors an altered POLD1. This, likewise, is consistent with human cancer studies that link C → A/G → T changes to POLD1 mutations (40). Second, the mutation rate varies greatly among the carcinomas, with tumor 5 having 907 genes significantly mutated, compared with 0 to 31 genes for tumors of similar or more advanced stages (Fig. 3A). This hypermutation is likely linked to defective DNA repair as well, because tumor 5 has as many as 24 DNA repair-associated genes mutated (Supplementary Table S3D). Third, many known human breast cancer genes are also mutated in these canine tumors (Supplementary Tables S3B and S3C), including BRCA1, IGFB2R, FOXC2, DLG2, and USH2A as described below. USH2A is one of the most significantly mutated genes in our study (P = 2.78E−12), having one nonsense-, 12 missense-, and three synonymous mutations (Fig. 3B; Supplementary Table S3F).

![Figure 3](https://example.com/fig3.png)

**Figure 3.** Coding sequence mutations are frequent in canine simple carcinomas; chromatin-modification genes are downregulated in canine complex carcinomas. A, the fractions (the y-axis) of somatic base substitution types of simple carcinomas (IDs indicated by the x-axis) detected by WES. The total number of significantly mutated genes in each tumor is also shown, and tumor 5 has many DNA repair genes mutated. B, synonymous (green dots) and nonsynonymous substitutions (yellow dots), and a nonsense mutation (red star) uncovered in the USH2A gene in tumor 5. C, the base substitution (compared with the dog reference genome) rates of the three sample types in coding regions with 30X to 300X RNA-seq read coverage. The P values were calculated by t tests. D, the heatmap of 751 genes differentially expressed at FDR < 0.2 between simple and complex carcinomas (red, upregulation; green, downregulation). Right, enriched functions of each gene cluster indicated; bottom, the 35 chromatin modifiers downregulated in complex carcinomas.
S3G). Critically, USH2A is also prominently mutated in human breast cancer, ranking as the 21st most significantly mutated gene in the Cancer Genome Atlas (TCGA) study (23). USH2A alterations may contribute to mammary cancer pathogenesis in both dogs and humans.

**Canine complex carcinomas have hardly any genomic CNAs and their sequence mutation rates also appear low**

Unlike simple carcinomas, we observed very few genomic CNAs in complex carcinomas, making their genomes appear normal (Fig. 2A; Supplementary Table S2A). Their sequence mutation rates are also low, based on our analysis with RNA-seq data. Briefly, to achieve a more accurate mutation finding, we used only regions with 30X to 300X RNA-seq read coverage, which distribute across the genome and amount to 4.5 to 9.4 Mb sequence per sample (Supplementary Table S3I). The analysis indicates that the mutation rates of complex carcinomas are significantly lower than those of simple carcinomas, but are comparable with those of normal mammary gland tissues (Fig. 3C).

**Numerous chromatin-modification genes are downregulated in canine complex carcinomas**

RNA-seq analysis revealed 751 total genes differentially expressed at FDR \( \leq 0.2 \) between simple and complex carcinomas (Fig. 3D). Strikingly, among these genes, chromatin modification and transcription regulation are the most significantly enriched functions (Supplementary Tables S4A–S4C). Indeed, a total of 35 known chromatin-modification genes were found to be downregulated in complex carcinomas (Fig. 3D; Supplementary Table S4E), and more than 40% of them remain so at \( P \leq 0.05 \) when further compared with normal mammary gland tissues. Moreover, chromatin modification stays as the most significantly enriched function amid genes (327 in total) differentially expressed among the three types of samples (Supplementary Fig. S3 and Supplementary Table S4H and S4I). The same overall conclusions were reached at FDR \(< 0.1 \) (Supplementary Table S4F and S4G).

Amid the 35 chromatin genes downregulated in complex carcinomas, 30 encode histone modifiers, covering methylation and demethylation, acetylation and deacetylation, and ubiquitination and deubiquitination (Fig. 3D). Intriguingly, both active and repressive modifiers were noted (see the paragraph that follows). Furthermore, the identified histone acetyltransferases and deacetylase modify histones H3, H4, and H2A, influencing not only gene transcription (e.g., CREBBP), but also chromatin packing (e.g., MSL1 on H4K16 acetylation; ref. 41). Besides histone-modification genes, other types of chromatin-remodeling genes were also found downregulated in complex carcinomas (Fig. 3D), most of which (e.g., ARID1B, ASFI1A, and DNMT3B) are known to be mutated in human cancers (42, 43).

To understand the significance of the observed change in chromatin-modification genes, many encoding histone modifiers (Fig. 3D), we investigated histone modification. Specifically, we performed IHC experiments to examine H3K9me3, a repressive modification that is associated with gene silencing and heterochromatin and for which six relevant genes are downregulated in complex carcinomas. These include H3K9 methyltransferase genes SETDB1, EHMT1, EHMT2, and SUZ12, along with the demethylase genes JMJD1C and PHF2 (Supplementary Table S4E). Meanwhile, we also examined H4-acetylation because at least eight of the downregulated genes (CREBBP, CSRP2BP, ING3, KAT2A, MSL1, MSL2, NCOA3, and SIRT1) are involved in histone acetylation or deacetylation. Another active modification, H3K4me3, was studied as well because the H3K4 methyltransferase genes SETD1A, ML1L2, and MLL4 are among those downregulated.

**In canine normal mammary glands, both active and repressive histone modifications are significantly depleted in myoepithelial cells when compared with luminal cells**

To understand the alteration in cancer, we first investigated canine normal mammary glands where both luminal and myoepithelial cells are clearly visible. These include the normal tissue from case 159, where myoepithelial cells form a nearly continuous layer surrounding the luminal cells (159N in Fig. 1A), and case 402421, where myoepithelial cells are not as prominent but are still noticeable (402421N in Fig. 1B). Interestingly, in these normal glands, active modifications, H4-acetylation and H3K4me3, and repressive modification H3K9me3 are both significantly depleted in the myoepithelial cells (Fig. 4A; Supplementary Fig. S4), with the intensity reduced by half in most cases (Fig. 4B), when compared with the luminal cells.

**In canine complex carcinomas, active modification H4-acetylation is abnormally enriched, whereas repressive modification H3K9me3 is aberrantly depleted**

Compared with normal mammary glands and simple carcinomas, complex carcinomas harbor significantly more myoepithelial cells (Fig. 1). Yet, unlike normal mammary glands (Fig. 4), both myoepithelial and luminal cells in complex carcinomas were found to be equally enriched with active modifications (Fig. 5A–H; Supplementary Fig. S5 and Supplementary Table S5). This is especially so for H4-acetylation, with the intensity being equal or stronger than luminal cells in normal mammary glands and in simple carcinomas. The repressive modification H3K9me3, to the contrary, becomes significantly more depleted in both cell types in complex carcinomas (Fig. 5I–L; Supplementary Fig. S5 and Supplementary Table S5).

**Redox genes are upregulated in canine ER\(^+\) carcinomas, whereas cell cycle and DNA repair genes are upregulated in canine ER\(^-\) carcinomas**

RNA-seq analyses also revealed a clear difference between canine ER\(^+\) and ER\(^-\) tumors (Fig. 6A), with most ER\(^+\) tumors being complex carcinomas, whereas most ER\(^-\) tumors being simple carcinomas. Among the 1,350 differentially expressed genes at FDR \( \leq 0.2 \) (Supplementary Table S6A), approximately half are upregulated in ER\(^+\) carcinomas and are significantly enriched in redox functions (Fig. 6B; Supplementary Table S6B, S6C, and S6E). These genes encode approximately 25 dehydrogenases or oxidases, and 32 gene...
products are associated with mitochondria, including the electron transport chain. Another half of the 1,350 differentially expressed genes are upregulated in ER+/C0 carcinomas, among which approximately 118 genes are associated with the cell cycle, for example, mitosis, spindle, microtubule cytoskeleton, etc. (Fig. 6B; Supplementary Table S6D and S6F). Other significant functions comprise DNA repair (38 genes) and protein serine/threonine kinase activity (17 genes). The same overall conclusions were reached at FDR/C20 (Supplementary Fig. S6A).

Canine simple carcinomas and the ER− complex carcinoma cluster closely with basal-like human breast carcinomas in PAM50 classification

To directly compare the canine mammary cancers with human breast cancers, we randomly selected 20 human tumors for each subtype among a total of 195 luminal A, 111 luminal B, 53 HER2-enriched, and 87 basal-like tumors of the TCGA RNA-seq study (23). This, along with all seven normal-like tumors in TCGA, amounts to 87 human carcinomas covering all five intrinsic subtypes. We then performed PAM50 clustering (44) on these 87 human carcinomas together with our 12 canine carcinomas. This analysis was repeated 100 times, ensuring that each TCGA tumor was sampled at least once. Notably, in 82 of 100 times, all canine simple carcinomas and the ER− complex carcinoma (ID 518) group with the human basal-like tumors. The remaining canine complex carcinomas (all ER+), however, fail to cluster with any specific human subtypes. One clustering example is shown in Fig. 6C and Supplementary Fig. S6B.

Discussion

In this study, we performed an initial comprehensive characterization of the genomes, transcriptomes, and epigenomes of two major canine mammary cancer histologic subtypes. Even with a small sample size (12 cases), the analysis reveals a remarkable molecular heterogeneity of spontaneous canine mammary cancers. It also emphasizes their unique value and raises a number of important questions that could profoundly affect human breast cancer research.

Canine simple carcinomas, without myoepithelial cell proliferation, harbor extensive genomic aberrations and are molecularly homologous to human breast carcinomas

Canine simple carcinomas investigated have no myoepithelial proliferation and are histologically comparable with human in situ or invasive ductal or lobular carcinomas.
Significantly, these canine cancers faithfully recapitulate key molecular features of human breast cancer. First, analogous to their human counterparts (18–23, 36), the genomes of these canine carcinomas harbor extensive genetic lesions, including numerous CNAs, fusion gene-creating translocation, equally complex superamplicon, and comparable sequence mutation types. The only exception is an inflammatory carcinoma, whose human equivalent (inflammatory breast cancer) is also devoid of CNAs (21). Second, notable human breast cancer genes (18–20, 23, 45, 46) are altered in these canine cancers as well. Examples include (i) amplification and/or overexpression of the oncogenes BRAF, MYC, PIK3CA, PIK3R1, CCND3, and TBX3; (ii) deletion and/or underexpression of tumor suppressors PTEN, PTPRD, and CDH1 (47); and (iii) mutations of BRCA1, NF1, MAP3K1, and RUNXI (Supplementary Table S7B). Third, many of the altered pathways are shared between the two species, for example, cell adhesion, Wnt signaling, PI3K signaling, and DNA repair (Fig. 7C; Supplementary Table S7; ref. 23), consistent with other canine mammary cancer studies (15, 17).

These strong molecular homologies make canine simple carcinomas valuable in human breast cancer research. For example, for cancers with large genomic CNAs, we can apply the dog–human comparison strategy for effective driver-passenger discrimination as described (7). Critically, as elegantly discussed in several publications (2–4), these canine cancers, which bridge a gap between traditional rodent models and human clinical trials, can significantly speed up new anticancer drug development. For example, for drugs targeting the PI3K pathway (Fig. 7C; ref. 48), their efficacy, toxicity, dosage, and schedule can be more accurately evaluated through clinical trials with canine patients, before entering human clinical trials. This will significantly reduce the cost and accelerate the drug discovery process.

Can canine simple carcinomas serve as a much-needed spontaneous cancer model of basal-like human breast cancer?

Canine simple carcinomas cluster with basal-like human tumors with an 82% chance in our PAM50 classification, indicating their closer resemblance to this subtype than other intrinsic subtypes. This may be explained by the observation that none of the canine tumors carry HER2 amplification or overexpression. Furthermore, many harbor extensive CNAs and are ER− with genes related to DNA repair and cell cycle significantly upregulated, consistent with the basal-like subtype profile (23). This is especially so considering that the ER− complex carcinoma clusters similarly as well. The only ER+ canine simple carcinoma has a prominent PTEN deletion, also a feature of basal-like tumors (23).

Clearly, studies with a larger sample size are needed to determine if canine simple carcinomas as a whole or even just a subset do indeed closely match the basal-like subtype. If confirmed, these canine cancers could serve as a much-needed...
spontaneous cancer model. Compared with other subtypes, basal-like cancers are aggressive, have a poor prognosis, and currently lack effective treatments. Canine mammary cancer could make significant contributions toward understanding and treating this worst subtype of human breast cancer.

Canine complex carcinomas, with myoepithelial cell proliferation, appear to originate from epigenomic rather than genomic alterations

Complex carcinomas, featuring dual proliferation of luminal and myoepithelial cells, likely originate from epigenomic, rather than genomic, abnormalities (Fig. 7). First, their genomes appear normal without CNAs detected and with sequence mutation rates as low as normal tissues. Thus, it is unlikely that these tumors arise from genetic aberrations, unlike simple carcinomas. Meanwhile, complex carcinomas could acquire genomic changes as they progress to later stages, as shown by tumor 518 (Supplementary Table S2A) and another complex carcinoma with pulmonary metastasis (14). Second, chromatin modification and transcription regulation stand out as the most enriched functions among genes differentially expressed between simple and complex carcinomas, with numerous chromatin-modification genes downregulated in complex carcinomas. Importantly, complex carcinomas are aberrantly enriched with the active histone modification H4-acetylation, whereas abnormally depleted with the

Figure 6. Subtype analysis reveals homology between canine mammary cancers and human breast cancers. A, examples of ER⁺ (ID 401188) and ER⁻ (ID 341400) canine carcinomas determined by IHC; scale bar, 100 μm. B, genes (1,350 total) differentially expressed between ER⁺ and ER⁻ canine carcinomas at FDR < 0.2. The tumor IDs are indicated. Tumor 401188 (purple) is ER⁺ but a simple carcinoma, whereas tumor 518 (orange) is ER⁻ but a complex carcinoma. Right, significantly enriched functions of each gene cluster indicated. C, canine simple carcinomas and the ER⁻ complex carcinoma cluster with the basal-like human breast carcinomas in PAM50 classification. The heatmap represents a clustering example of 12 canine tumors and 87 TCGA human tumors (see text). The composition of each cluster specified at the top of the heatmap is explained at the right side.
A repressive modification H3K9me3. Thus, it is possible that the epigenomes of complex carcinomas are altered, turning on genes that normally should be silenced to promote tumorigenesis. Obviously, more studies are needed to confirm this possibility and to understand the underlying mechanisms.

Myoepithelial cell proliferation is rare in human breast cancer (32, 33). As a result, myoepithelial cells receive far less attention than luminal cells and are poorly understood (25, 27, 28, 30). However, the few laboratories that study myoepithelial cells have noted their importance. For example, myoepithelial cells are thought to be a part of the mammary stem cell niche, mediate the cross-talk between luminal cells and extracellular matrix, contribute to the maintenance of the apicobasal polarity of luminal cells, and serve as a tumor suppressor (26, 29, 31). Canine mammary cancer, in which myoepithelial cell proliferation is much more common, provides an ideal system to address such functions and to better understand the second major cell lineage of the mammary gland (e.g., by answering questions such as whether a prolonged luteal phase promotes myoepithelial cell proliferation).

### Do canine complex carcinomas derive from mammary gland stem cells or luminal/myoepithelial common progenitors?

Several observations indicate the possibility that complex carcinomas arise from mammary gland stem cells or luminal/myoepithelial common progenitors (Fig. 7A). First, one of these tumors (ID 518) expresses the pluripotency marker SOX2, indicating stem cell property. Second, unlike normal mammary glands that present a clearly different epigenomic landscape between luminal and myoepithelial cells, no such difference was observed in complex carcinomas. This indicates that proliferating luminal and myoepithelial cells in complex carcinomas may have derived from altered common progenitors. Third, compared with simple carcinomas and normal mammary tissues, glucose metabolic genes are upregulated in complex carcinomas, consistent with this stem cell or progenitor origin theory. For simple carcinomas, we hypothesize that they originate from either luminal progenitors, because of their close resemblance to the basal-like subtype, which is reported to have a luminal progenitor origin (49), or differentiated luminal cells because of luminal cell properties (see case 159 in Fig. 1). Of course, further studies with a larger sample size are needed to test these hypotheses.

In summary, we performed the first comprehensive characterization of the genomes, transcriptomes, and epigenomes of canine simple carcinomas and complex carcinomas, two major histologic subtypes of canine mammary cancer. The analysis reveals that canine simple carcinomas, which have no myoepithelial cell proliferation and histologically match typical human breast carcinomas, faithfully recapitulate many molecular features of human breast cancer. Notably, canine simple carcinomas closely cluster with basal-like human breast tumors in PAM50 classification, and, thus, could serve as a much-needed spontaneous cancer model for the basal-like subtype. Canine complex carcinomas are characterized with dual proliferation of luminal and myoepithelial cells, which is rare in human breast cancer. Our analysis indicates that these canine cancers may arise from epigenomic rather genomic alterations. Canine complex carcinomas, hence, provide a unique system to investigate the roles of myoepithelial cells, the second major cell lineage of the mammary gland, in normal developmental and pathogenic processes.
Data access
Sequence data have been submitted to the NCBI SRA database with accession numbers SRP023115, SRP023472, and SRP024250. aCGH data have been submitted to the GEO database with the accession number GSE54535.

Disclosure of Potential Conflicts of Interest
No potential conflicts of interest were disclosed.

Authors’ Contributions
Conception and design: D. Liu, S. Zhao
Development of methodology: D. Liu, H. Xiong, N.C. Northrup, S. Zhao
Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): D. Liu, A.E. Ellis, N.C. Northrup, C.O. Rodriguez Jr, S. Dalton, S. Zhao
Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): D. Liu, S. Zhao
Writing, review, and/or revision of the manuscript: D. Liu, A.E. Ellis, S. Zhao
Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): D. Liu, H. Xiong, A.E. Ellis
Study supervision: R.M. O’Regan

References

Acknowledgments
The authors thank Drs. Timothy D. Read, Shawn Levy, Steven H. Miller, and their respective teams, and the BGI, for their outstanding sequencing or aCGH work; Teri Guerrero, Irene Mok, and Dr. Susan E. Lana for their help on sample collection; Dr. John M. Rosenfield of Millipore for providing antibodies; Drs. Nancy Manley and Jie Li for their help on IHC and H&E analyses; and Drs. Dong M. Shin, J. David Puett, Charles M. Perou, and Malcolm Hayes for their help on the study. Confocal imaging was performed at the UGA Biomedical Microscopy Core.

Grant Support
This work was funded by the American Cancer Society, the Georgia Cancer Coalition, NCI R01 CA182093, and the AKC Canine Health Foundation (to S. Zhao), as well as by pilot project funds from NCI P50 CA128613 (PI, Dr. Dong M. Shin) and GM065354 (PI, S. Dalton). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked advertisement in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received February 11, 2014; revised June 24, 2014; accepted June 29, 2014; published OnlineFirst July 31, 2014.

Canine and Human Mammary Cancer Molecular Comparison

Other (consultation on issues related to animal cancer, references, sample acquisition): N.C. Northrup

Published OnlineFirst July 31, 2014; DOI: 10.1158/0008-5472.CAN-14-0392

Downloaded from cancerres.aacrjournals.org on September 15, 2017. © 2014 American Association for Cancer Research.
Molecular Homology and Difference between Spontaneous Canine Mammary Cancer and Human Breast Cancer

Deli Liu, Huan Xiong, Angela E. Ellis, et al.


Updated version
Access the most recent version of this article at:
doi:10.1158/0008-5472.CAN-14-0392

Supplementary Material
Access the most recent supplemental material at:
http://cancerres.aacrjournals.org/content/suppl/2014/08/01/0008-5472.CAN-14-0392.DC1

Cited articles
This article cites 49 articles, 8 of which you can access for free at:
http://cancerres.aacrjournals.org/content/74/18/5045.full#ref-list-1

Citing articles
This article has been cited by 1 HighWire-hosted articles. Access the articles at:
http://cancerres.aacrjournals.org/content/74/18/5045.full#related-urls

E-mail alerts
Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions
To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions
To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.