CEACAM Gene Family Mutations Associated With Inherited Breast Cancer Risk – A Comparative Oncology Approach to Discovery

Anna L. W. Huskey1,2, Isaac McNeely1 and Nancy D. Merner1*

1 Department of Pathobiology, College of Veterinary Medicine, Auburn University, Auburn, AL, United States, 2 Department of Drug Discovery and Development, Harrison School of Pharmacy, Auburn University, Auburn, AL, United States

Introduction: Recent studies comparing canine mammary tumors (CMTs) and human breast cancers have revealed remarkable tumor similarities, identifying shared expression profiles and acquired mutations. CMTs can also provide a model of inherited breast cancer susceptibility in humans; thus, we investigated breed-specific whole genome sequencing (WGS) data in search for novel CMT risk factors that could subsequently explain inherited breast cancer risk in humans.

Methods: WGS was carried out on five CMT-affected Gold Retrievers from a large pedigree of 18 CMT-affected dogs. Protein truncating variants (PTVs) detected in all five samples (within human orthlogs) were validated and then genotyped in the 13 remaining CMT-affected Golden Retrievers. Allele frequencies were compared to canine controls. Subsequently, human blood-derived exomes from The Cancer Genome Atlas breast cancer cases were analyzed and allele frequencies were compared to Exome Variant Server ethnic-matched controls.

Results: Carcinoembryonic Antigen-related Cell Adhesion Molecule 24 (CEACAM24) c.247dupG:p.(Val83Glyfs*48) was the only validated variant and had a frequency of 66.7% amongst the 18 Golden Retrievers with CMT. This was significant compared to the European Variation Archive (p-value 1.52 × 10^{-8}) and non-Golden Retriever American Kennel Club breeds (p-value 2.48 × 10^{-5}). With no direct ortholog of CEACAM24 in humans but high homology to all CEACAM gene family proteins, all human CEACAM genes were investigated for PTVs. A total of six and sixteen rare PTVs were identified in African and European American breast cancer cases, respectively. Single variant assessment revealed five PTVs associated with breast cancer risk. Gene-based aggregation analyses revealed that rare PTVs in CEACAM6, CEACAM7, and CEACAM8 are associated with European American breast cancer risk, and rare PTVs in CEACAM7 are associated with breast cancer risk in African Americans. Ultimately, rare PTVs in the entire CEACAM gene family are associated with breast cancer risk in both European and African Americans with respective p-values of 1.75 × 10^{-13} and 1.87 × 10^{-04}. 

*Correspondence: Nancy D. Merner
ndm0011@auburn.edu

Specialty section:
This article was submitted to Cancer Genetics and Oncogenomics, a section of the journal Frontiers in Genetics

Received: 30 April 2021
Accepted: 05 July 2021
Published: 10 August 2021

Citation: Huskey ALW, McNeely I and Merner ND (2021) CEACAM Gene Family Mutations Associated With Inherited Breast Cancer Risk – A Comparative Oncology Approach to Discovery. Front. Genet. 12:702889.
doi: 10.3389/fgene.2021.702889
INTRODUCTION

Breast cancer is a serious health concern. Amongst both sexes, it globally ranks as the second most commonly diagnosed type of cancer and the second leading cause of cancer-related deaths, accounting for ~2.1 million new diagnoses and 626,679 deaths in 2018 (Bray et al., 2018). Worldwide, it is also the most common cancer diagnosed in women and the overall leading cause of cancer-related female deaths (Bray et al., 2018). In the United States, 2020 estimates predicted breast cancer to be the leading site of new cancer diagnoses in women and the second leading cause of cancer-related deaths, resulting in 276,480 new diagnoses and 42,170 deaths (American Cancer Society, 2020). Advances in breast cancer research have translated to better disease screening, diagnosis, and treatment, but new research questions continuously arise as time and medical needs progress (Cardoso et al., 2017).

Comparative oncology, which is the study of cancer biology and therapy in spontaneous, naturally-occurring cancers in companion animals, provides valuable models of human cancer that have and will continue to make research advances (Garden et al., 2018). Recent studies comparing canine mammary tumors (CMTs) and human breast cancers have revealed notable tumor similarities, identifying shared expression profiles and acquired mutations (Liu et al., 2014; Ettlin et al., 2017; Lee et al., 2018, 2019; Kim et al., 2019; Gray et al., 2020). CMTs can also provide a model of hereditary breast cancer susceptibility in humans, especially considering similar genetics and familial clustering (Goebel and Merner, 2017; Guo et al., 2014). While most CMT studies investigating inherited risk have focused on identifying genetic variants in orthologs of known human breast cancer risk genes (Goebel and Merner, 2017; Huskey et al., 2020), in this study, we investigate breed-specific whole genome sequencing (WGS) data in search for novel CMT risk factors. WGS studies have been used to make numerous disease gene discoveries in dogs, many of which clearly translated to human health (Spina et al., 2006; Beauchemin and Arabzadeh, 2013; Han et al., 2020). Aberrant expression of many CEACAM genes have previously been associated with tumorigenesis, and CEACAM gene products are recognized as clinically-relevant tumor markers (Kuespert et al., 2006; Beauchemin and Arabzadeh, 2013; Han et al., 2020). This is the first association to be reported between CEACAM gene mutations and inherited cancer risk.

MATERIALS AND METHODS

**Golden Retriever Pedigree and WGS**

As previously described by Huskey et al. (2020), blood- or buccal-derived DNA samples were obtained from 18 CMT-affected Golden Retrievers from the Canine Health Information Center (CHIC) DNA repository, and a pedigree was constructed linking all 18 dogs in one large pedigree. Five of those Golden Retrievers (three females and two males) were selected for WGS. This number was influenced by the cost of WGS. Furthermore, aiming to identify breed-specific mutations, distantly related dogs were selected, including two males since male breast cancer is associated with hereditary disease (Huskey et al., 2020). The WGS data was processed through a bioinformatics pipeline (Huskey et al., 2020). Upon alignment to the CanFam3.1 reference genome and annotation using gene predictions from Ensembl build version 75, a script was written to isolate PTVs found in all five Golden Retriever samples. PTVs were defined as single nucleotide variants (SNVs) that resulted in a premature stop codon or abrogated a splice site, and small insertions or deletions (indels) that changed a transcript’s reading frame. Upon filtering, the genes with PTVs were classified into two different groups, orthologs of human genes or non-orthologs. Polymerase chain reaction (PCR) and Sanger sequencing were carried out to validate the PTVs in human orthologs. CEACAM24 c.247dupG>Gpc.(Val83Glyfs*48) was the only validated variant. Following validation, the 13 remaining CMT-affected Golden Retrievers underwent PCR and Sanger sequencing to determine their mutation status.

**Canine Controls**

As a convenient, publically available, online canine genetic variant repository, the European Variation Archive1 was initially used to note the allele frequency of CEACAM24 c.247dupG>Gpc.(Val83Glyfs*48). The European Variation Archive

---

1https://www.ebi.ac.uk/eva/?eva-study=PRJEB24066
provides high quality WGS variant calls of over 200 dogs from multiple breeds (breed and sex information was unknown). The data was obtained through Ensembl by accessing the canine gene’s “Variant table” under “Genetic Variation”; for a particular variant, “Population genetics” information was given, including European Variation Archive allele frequencies (Zerbino et al., 2018). Furthermore, additional splicing, frame-shifting, and stop gain mutations within the other dog CEACAM genes were investigated through Ensembl transcripts (CEACAM16: ENSCAF00000044174; CEACAM18: ENSCAF0000004587; CEACAM20: ENSCAF00000047731; CEACAM24: ENSCAF00000047960; CEACAM28:
ENSCAFT00000022623). CEACAM1, CEACAM23, and CEACAM30 did not have variant information available in Ensembl for European Variation Archive data.

Through the CHIC repository, blood or buccal-swab derived DNA from purebred, American Kennel Club registered dogs were randomly selected and obtained to determine the frequency of CEACAM24 c.247dupG;p.(Val83Glyfs*48). This included DNA from Golden Retrievers (n = 87), as well as 13 other breeds, including Petit Basset des Griffon (n = 10), Gordon Setter (n = 8), Australian Cattle Dog (n = 10), Siberian Husky (n = 10), Dalmatian (n = 10), Irish Setter (n = 9), Welsh Pembroke Corgi (n = 10), Standard Schnauzer (n = 10), Newfoundland (n = 10), Keeshond (n = 10), Great Dane (n = 8), Doberman Pinscher (n = 10), and Boxer (n = 10). PCR and Sanger sequencing were carried out to determine CEACAM24 c.247dupG;p.(Val83Glyfs*48) genotypes of each dog.

Canine Statistical Analyses
Upon determining CEACAM24 c.247dupG;p.(Val83Glyfs*48) allele frequencies, p-values were generated using the Fisher’s Exact Test in R (v 3.5.1), comparing allele differences in Golden Retriever to control dogs, including both European Variation Archive and CHIC DNA samples.

Dog and Human CEACAM Protein Analyses
EMBOSS water alignment (Madeira et al., 2019) was carried out to determine the level of homogeneity between the dog CEACAM24 protein and other dog and human CEACAM proteins. Additionally, InterPro (Hunter et al., 2009) and the Eukaryotic Linear Motif (ELM) resource (Kumar et al., 2020) were used to identify CEACAM domains and binding motifs, respectively.

Human CEACAM Gene Analysis – The Cancer Genome Atlas
Due to the homogeneity of the CEACAM gene family and no direct ortholog of dog CEACAM24 in humans, all human CEACAM family genes were investigated for rare PTVs in The Cancer Genome Atlas (TCGA) breast cancer cohort. Investigating inherited risk, only blood-derived exomes of breast cancer cases were analyzed. Overall, whole-exome binary sequence alignment mapping (BAM) files were downloaded using the Genomic Data Commons (GDC) Data Portal Repository through approved research project #10805. To acquire the samples, the specific filters under the “Cases” category included: Project (TCGA-BRCA), Samples Sample Type (Blood Derived Normal), and Race (“Black or African American” and “White”). The samples were further filtered under the “Files” category, including Experimental Strategy (WXS) and Data Format (BAM). A total of 170 sample files were obtained for African Americans and 650 for European Americans. These files were downloaded using the GDC Data Transfer Tool (version 1.2.0). Only individuals with known ages of breast cancer onset were used in this study; as a result, one European American and two African American BAM

| Data set/Cohort | Dog breed | # of dogs | # of HOM | # of HET | Minor allele frequency | p-value for comparison to CMT affected Golden Cohort |
|----------------|-----------|-----------|----------|----------|------------------------|------------------------------------------------------|
| CMT Affected   | Golden Retriever | 18   | 6    | 9 | 66.7 | - |
| CHIC United States Breed Specific Controls | Golden Retriever | 87 | 42 | 34 | 67.8 | 0.3334 |
| CHIC United States Non-Golden Retriever Controls | Petit Basset Griffon Vendeen | 10 | 7 | 2 | 80.0 | 2.48 x 10⁻⁵ |
| Gordon Setter | 8    | 5    | 2    | 75.0 | | |
| Australian Cattle Dog | 10   | 4    | 2    | 50.0 | | |
| Siberian Husky | 10   | 4    | 1    | 45.0 | | |
| Dalmatian | 10   | 3    | 2    | 40.0 | | |
| Irish Setter | 9    | 0    | 1    | 5.6 | | |
| Welsh Pembroke Corgi | 10   | 0    | 0    | 0.0 | | |
| Standard Schnauzer | 10   | 0    | 0    | 0.0 | | |
| Newfoundland | 10   | 0    | 0    | 0.0 | | |
| Keeshond | 10   | 0    | 0    | 0.0 | | |
| Great Dane | 8    | 0    | 0    | 0.0 | | |
| Doberman Pinscher | 10   | 0    | 0    | 0.0 | | |
| Boxer | 10   | 0    | 0    | 0.0 | | |
| Totals and Avg MAF of CHIC Non-Golden Retriever Controls | 125 | 23 | 10 | 22.4 | | |
| European Variation Archive Controls | European General Dog Population | 196 | 12 | 44 | 17.3 | 1.52 x 10⁻⁶ |

The bold values represent significant p-values, p-values less than 0.05.
files were removed from further bioinformatics processing and statistical analysis.

The downloaded BAM files, which had previously been aligned to the hg38 human reference genome, were processed using the remaining steps of a pipeline adapted from the Genome Analysis Toolkit’s (GATK’s) best practices pipeline (Van der Auwera et al., 2013). Base quality scores were recalibrated using BaseRecalibrator and then HaplotypeCaller was used to generate variant calling format (gVCF) files (GATK version 3.6). GenotypeGVCFs was used to merge the individual gVCF files based on ethnicity (GATK version 3.6). The European American files were merged in batches of approximately 200 using GATK’s (version 3.6) CombineGVCFs prior to merging into a single VCF file with GenotypeGVCFs. The two ethnic specific VCF files were then processed through a variant quality score recalibration using VariantRecalibrator (GATK version 3.6), and, as recommended, SNVs were filtered using a pass filter of 99.5%, and indels were filtered using a slightly lower pass filter of 99.0% (Van der Auwera et al., 2013). Variants in CEACAM1 (NM_001184815; chr19:42507306-42528481), CEACAM3 (NM_001815 at chr19:41796587-41811554), CEACAM4 (NM_001817; chr19:41618971-41627074), CEACAM5 (NM_004363; chr19:41708626-41730421), CEACAM6 (NM_001816 at chr19:41755530-41772210), CEACAM7 (NM_006890; chr19:41673303-41688270), CEACAM8 (NM_001039213 at chr19:42580243-42594924), CEACAM16 (NM_001039213 at chr19:44699151-44710718), CEACAM18 (NM_001278392; chr19:51478643-51490605), CEACAM19 (NM_064604; chr19:44671452-44684355), CEACAM20 (NM_001102597; chr19:44506159-44529675), and CEACAM21 (NM_001098506; chr19:41576166-41586844) were then extracted from the ethnic specific VCF files and annotated using ANNOVAR (version June2017). Variants were filtered to include rare PTVs with ethnic-specific minor allele frequencies of <1% in Exome Variant Server (EVS; National Heart, Lung, and Blood Institute (NHLBI) Exome Sequencing Project) (Exome Variant Server, 2019).

**Human Statistical Analyses**

Using the Fisher’s exact test (Sprent, 2011) in R (v 3.5.1), individual PTVs were assessed to compare allele frequency differences between ethnic-specific TCGA breast cancer cases and EVS controls. The Fisher’s method was used for gene-based and gene family-based aggregation analyses (Fisher, 1925; Sutton et al., 2000). The R tool “sumlog” (in the “metap” package) was used to combine p-values for each aggregation test. To accommodate for the one-sided nature of the Fisher exact test p-values, compliments of p-values in the opposite direction were used in the calculations for the Fisher’s method aggregation analyses.

**Human Mutation Analysis**

Mutalyzer was used to determine the effect of frame-shifting and non-sense variants on the coded protein (Wildeman et al., 2008). Human splicing mutations that affected non-protein-coding exons of the mRNA, specifically in the 3′ untranslated region (UTR), were analyzed using the miRDB tool to identify microRNA binding sites potentially lost due to a splicing mutation (Chen and Wang, 2020). For each gene harboring a splice mutation affecting non-protein-coding exons, microRNA binding sites within the 3′ UTR with a target score of ≥80 were

| TABLE 2 | Homology of dog and human ceacam proteins to dog CEACAM24 protein. |
| Species | Gene name | Protein accession | % Identity | % Similarity |
|---------|-----------|-------------------|------------|-------------|
| Dog     | CEACAM1   | NP_00101026       | 52.2       | 58.4        |
|         | CEACAM16  | ENSCAGP00000039084| 22.5       | 37.7        |
|         | CEACAM18  | ENSCAGP00000058450| 19.3       | 32.5        |
|         | CEACAM20  | ENSCAGP00000036293| 21.2       | 31.9        |
|         | CEACAM23  | NP_001091021      | 38.4       | 40.8        |
|         | CEACAM24  | NP_001091023      | 100        | 100         |
|         | CEACAM28  | NP_001091015      | 42.2       | 46.3        |
|         | CEACAM30  | NP_001091022      | 53.6       | 58.3        |
|         | Average of all Dog CEACAM proteins compared to Dog CEACAM24 (excluding CEACAM24 from analysis) | 35.6 | 43.7 |
| Human   | CEACAM1   | NP_001171744      | 53.1       | 60.8        |
|         | CEACAM3   | NP_001806         | 47         | 58.2        |
|         | CEACAM4   | NP_001808         | 50.4       | 63.4        |
|         | CEACAM5   | NP_004354         | 53.2       | 61          |
|         | CEACAM6   | NP_002474         | 37.8       | 48          |
|         | CEACAM7   | NP_008821         | 45.1       | 58.3        |
|         | CEACAM8   | NP_001807         | 53.8       | 63.6        |
|         | CEACAM16  | NP_00134302       | 28         | 43.5        |
|         | CEACAM18  | NP_001265321      | 26.9       | 46.2        |
|         | CEACAM19  | NP_0064604        | 23.7       | 38.1        |
|         | CEACAM20  | NP_001090607      | 25.7       | 39.9        |
|         | CEACAM21  | NP_001091976      | 34.1       | 42.3        |
|         | Average of all Human CEACAM proteins compared to Dog CEACAM24 | 39.9 | 51.9 |
noted. The top five ranked microRNA targets were investigated for previous cancer (specifically, hereditary breast and ovarian cancer (HBOC) syndrome) associations.

RESULTS

Upon filtering the WGS data, 12 different PTVs were detected in all five Golden Retrievers, four of which were within human orthologs. Only one PTV, a frame-shifting mutation in *CEACAM24* (c.247dupG:p.(Val83Glyfs*48)) was determined to be a true positive upon validation (Figure 1). This mutation had a frequency of 66.7% amongst the 18 Golden Retrievers with CMT in this study (Table 1). Upon comparing that frequency to the 17.3% allele frequency in the European Variation Archive, a p-value of $1.52 \times 10^{-8}$ was generated. Representing dogs from another continent and not knowing the breeds of the European Variation Archive, the frequency of *CEACAM24* c.247dupG:p.(Val83Glyfs*48) was subsequently determined in different American Kennel Club

![Figure 2](image-url)
breeds (Table 1). There was no statistically significant difference between Golden Retriever CMT cases and controls. However, there was a significant difference between Golden Retrievers cases and other American Kennel Club breeds ($2.48 \times 10^{-5}$; Table 1). The CEACAM24 c.247dupGp.(Val83Glyfs*48) allele frequency ranged from 0 to 80% in the assessed breeds (Table 1). CEACAM24 c.247dupGp.(Val83Glyfs*48) abolishes the extracellular region, the transmembrane domain, and part of the cytoplasmic region, including the Ig V-set domain (Figures 1C,D).

Homology analysis revealed that the dog CEACAM proteins were, on average, 43.7% similar to the dog CEACAM24 protein (Table 2 and Figure 2A). Similarly, there were many related functional domains and high homology between the dog CEACAM24 protein and the human CEACAM proteins, averaging 51.9% similarity (Table 2 and Figure 2). This homology, along with the fact that there is no direct human ortholog of dog CEACAM24, prompted all human CEACAM genes (Figure 2B) to be investigated for rare PTVs in the TCGA breast cancer cohort.

A total of six rare PTVs were identified in African Americans and sixteen in European Americans breast cancer cases (Supplementary Tables 1, 2). Single variant assessment revealed five variants associated with breast cancer risk, three of which were associated each with European and African American breast cancer (Table 3 and Figures 3, 4). One variant, CEACAM7 c.195C > A:p.(Y65X), was associated with breast cancer risk in both ethnicities (Table 3 and Figure 3). Two stop gain mutations in CEACAM4 were associated with African American breast cancer (Table 3 and Figure 3), and two splicing mutations were associated with European American breast cancer, one in CEACAM6 and another within CEACAM8 (Table 3 and Figure 4). Both of those splicing mutations affect non-protein-coding exons in the 3′ UTR, which, instead of truncating the protein, potentially disrupt key microRNA binding sites previously associated with cancer (Table 4 and Figure 4). Overall, gene-based aggregation analyses revealed that rare PTVs in CEACAM6, CEACAM7, and CEACAM8 are associated with European American breast cancer risk, and rare PTVs in CEACAM7 are associated with breast cancer risk in African Americans (Table 5). Ultimately, rare PTVs in the entire CEACAM gene family are associated with breast cancer risk in both European and African Americans with respective $p$-values of $1.75 \times 10^{-13}$ and $1.87 \times 10^{-04}$ (Table 5).

### DISCUSSION

Utilizing a comparative oncology approach, our team identified CEACAM24 c.247dupGp.(Val83Glyfs*48) in Golden Retrievers with CMT and subsequently determined that rare PTVs in the entire CEACAM gene family were associated with inherited breast cancer risk in humans. We previously described a large Golden Retriever pedigree with segregating CMT, carried out WGS on five selected Golden Retriever cases, and highlighted variants in orthologs of human breast cancer susceptibility genes (Huskey et al., 2020). In this current study, we used

| Gene name | Protein variant name | rs ID EA AA | AA MAF (%) Mutation specific $p$-values | EA MAF (%) Mutation specific $p$-values | TCGA EA MAF (%) Mutation specific $p$-values | EVS AA MAF (%) Mutation specific $p$-values | TCGA EA MAF (%) Mutation specific $p$-values |
|------------|----------------------|------------|-----------------------------------------------|-----------------------------------------------|-----------------------------------------------|-----------------------------------------------|-----------------------------------------------|
| CEACAM4: NM_001817 | stopgain | c.367C>T | rs147663846 - - | 0.20 0.89 | 0.04803 |
| CEACAM6: NM_002483 | splicing | c.*40 +2T | rs782698255 0.00 0.46 | 7.40E-06 |
| CEACAM7: NM_006890 | stopgain | c.195C>A | rs782316651 0.00 10.79 | 2.20E-16 |
| CEACAM8: NM_001816 | splicing | c.*40 +2T | rs748512513 0.00 0.00 | 2.20E-06 |

The bold values represent significant $p$-values, $p$-values less than 0.05.
the same WGS dataset to identify novel variants that could be influencing Golden Retriever CMT susceptibility. We isolated PTVs found in all five sequenced Golden Retriever samples, and, upon validation, determined the mutation status in the 13 remaining CMT-affected Golden Retrievers within the pedigree. CEACAM24 c.247dupG;p.(Val83Glyfs*48) was the only validated variant and had an allele frequency of 66.7% amongst the 18 CMT-affected dogs. Despite not being recognized as a breed highly affected by CMT, Golden Retrievers have a higher prevalence of cancer compared to many dog breeds with 65% of Golden Retrievers in the United States succumbing to the disease (Dobson, 2013; Salas et al., 2015; Kent et al., 2018). The Golden Retriever CEACAM24 c.247dupG;p.(Val83Glyfs*48) allele frequency and cancer mortality rate are very similar.

The CMT-affected Golden Retrievers within this study can all be linked back to a sire in the United States from the 1950s, which was shortly after the registration of the breed with the American Kennel Club. Since importation to and registration in the United States, Golden Retrievers in Europe and the United States are considered two distinct populations, as breeding between the two continents is rare and unique gene pools have been established due to strict breeding standards and the popular-sire effect (Brackman, 2020). Cancer mortality in European-bred Golden Retrievers has been reported to be 38.8%, which is much lower than Golden Retrievers in the United States (65%) (Dobson, 2013; Kent et al., 2018). These differences could be explained by distinct genetic risk factors. The allele frequency of CEACAM24 c.247dupG;p.(Val83Glyfs*48) in the European Variant Archive was 17.3%, which corresponded to a p-value of $1.52 \times 10^{-8}$ when compared to our CMT-affected Golden Retrievers from the United States. However, in addition to not knowing breed-specific information in
the European Variant Archive, genetic bottlenecks upon importation to the United States need to be acknowledged. Thus, comparing allele frequencies to a United States dog population with known breed status was important, which can be determined through American Kennel Club registration. Overall, CEACAM24 c.247dupG:p.(Val83Glyfs*48) appears to be common in Golden Retrievers in the United States with an allele frequency of 67.8%, which is not significantly different from the CMT-affected Golden Retriever cases. However, that allele frequency was determined by screening 87 Golden Retrievers from the CHIC repository with unknown disease diagnoses and age at sample submission. This is not ideal for canine cancer studies; older dogs (> 8 years of age) with unaffected CMT-status are recommended (Tonomura et al., 2015; Hayward et al., 2016). In saying that, if CEACAM24 c.247dupG;p.(Val83Glyfs*48) truly is a high-frequency allele in Golden Retrievers due to a genetic bottleneck in the United States, it can explain why 65% of Golden Retrievers succumb to cancer (Kent et al., 2018).

Regarding the assessment of other American Kennel Club breeds, an overall CEACAM24 c.247dupG;p.(Val83Glyfs*48) allele frequency of 22.4% was revealed, which was significantly different from CMT-affected Golden Retriever cases. Noting the small sample sizes of each breed, over half of the assessed breeds showed no presence of the variant. However, some breeds contained the variant at higher levels; most notably, Petit Basset Griffon Vendeen, Gordon Setter, Australian Cattle Dog, Siberian Husky, and Dalmatian. Petit Basset Griffon Vendeen, which had the highest allele frequency, has a cancer mortality rate of 33% (Dobson, 2013). In a United Kingdom study, Dalmatians, Gordon Setters, and Siberian Huskies were found to have cancer mortality rates ranging from 19.1 to 31.8% (Dobson, 2013), and Australian Cattle Dogs have a rate of 27% (Petmed, 2014).

CEACAM24 is a part of the dog CEACAM gene family (Figure 2A), which is a subdivision of the immunoglobulin superfamily of cell adhesion molecules (IgCAMs) (Smith and Xue, 1997; Kuespert et al., 2006). All IgCAMs, and hence all CEACAM proteins, are characterized by having at least one immunoglobulin (Ig)-like domain (Figure 2). CEACAM genes have diverse functions in both dogs and humans, including cell-cell adhesion, cell signaling, immunity/inflammation, angiogenesis, and tumor development, progression and metastasis (Kuespert et al., 2006; Kammerer et al., 2007; Kammerer and Zimmermann, 2010; Beauchemin and Arabzadeh, 2013; Han et al., 2020). CEACAM24 c.247dupG;p.(Val83Glyfs*48) abolishes the extracellular region, the transmembrane domain, and part of the cytoplasmic region, including the Ig V-set domain; thus, it is presumed to be a loss-of-function mutation. According to Ensembl, no other stop gain or frame-shifting variants have been identified in dog CEACAM genes. However, one splicing mutation in CEACAM28 (c.1415-2A > G) was identified, which had a 34% allele frequency within the European Variation Archive. The CEACAM gene family is present in many mammalian species but has evolved in a highly species-specific manner, heavily influenced by pathogen/host coevolution (Kammerer et al., 2007; Kammerer and Zimmermann, 2010; Weichselbaumer et al., 2011). Despite phylogenetic discordance of dog and human CEACAM genes (Weichselbaumer et al., 2011), our analyses revealed there is high homology between the dog CEACAM24 protein and the human CEACAM proteins, averaging 51.9% similarity. This homology, along with the fact that there is no direct human ortholog of the CEACAM24 gene, prompted all human CEACAM genes to be investigated for rare PTVs in the TCGA breast cancer cohort.

There are 12 human CEACAM genes, all of which cluster on chromosome 19q13.2-19q13.4. Over the years, genetic markers in that region have been associated with many different types of cancer susceptibility, including breast cancer (Rockenbauer et al., 2002; Yin et al., 2002; Nexo et al., 2003, 2008; Vogel et al., 2004; Amin Al Olama et al., 2013; Gao et al., 2018).
TABLE 4 | Top five miRNA binding sites for both CEACAM6 and CEACAM8 and previous cancer associations.

| Gene target name | miRNA name | Previous cancer association | Previous HBOC association |
|------------------|------------|----------------------------|--------------------------|
| CEACAM6          | mR-3119    | Yes                        | No                       |
|                  | mR-766-3p  | Yes                        | Yes                      |
|                  | mR-6512-3p | Yes                        | Yes                      |
|                  | mR-6720-5p | Yes                        | Yes                      |
|                  | mR-5702    | Yes                        | Yes                      |
| CEACAM8          | mR-661     | Yes                        | Yes                      |
|                  | mR-9903    | Yes                        | Yes                      |
|                  | mR-616-5p  | Yes                        | Yes                      |
|                  | mR-371b-5p | Yes                        | Yes                      |
|                  | mR-4635    | Yes                        | Yes                      |

Nonetheless, inherited mutations in CEACAM genes have yet to be associated with inherited risk of cancer (Zheng et al., 2011; Kammerer et al., 2012; Wang et al., 2015). aberrant expression of many CEACAM genes have been associated with tumorigenesis, and CEACAM gene products are recognized as clinically-relevant tumor markers (Kuespert et al., 2006; Beauchemin and Arabzadeh, 2013; Han et al., 2020). Regarding breast cancer, CEACAM1 has been shown to be down-regulated compared to normal breast tissue (Yang et al., 2015), similar to its expression in prostate (Busch et al., 2002; Liu J. et al., 2020), endometrial (Bamberger et al., 1998), gastric (Takeuchi et al., 2019) and colon cancer (Fournes et al., 2001; Song et al., 2011), identifying it as a tumor suppressor. It has also been demonstrated that CEACAM5 (Iqbal et al., 2017; Powell et al., 2018), CEACAM6 (Marqa et al., 2008; Tsang et al., 2013; Iqbal et al., 2017; Rizeq et al., 2018), and CEACAM19 (Michaelidou et al., 2013; Estiar et al., 2017) are overexpressed in breast cancer and are associated with enhanced tumor invasiveness and metastasis. Conversely, CEACAM6 and CEACAM8 co-expression inhibits proliferation and invasiveness of breast cancer cells (Iwabuchi et al., 2019). Additionally, CEACAM gene splice variants have been suggested to play a role in breast cancer tumorigenesis (Gaur et al., 2008; Zisi et al., 2020). Lastly, through exome sequencing, Li et al. observed loss of heterozygosity of CEACAM1, CEACAM3, CEACAM5, CEACAM6, CEACAM7, and CEACAM8 in breast cancer tumors that were associated with metastasis, suggesting that this closely-linked gene family regulates tumorigenesis and metastasis synergistically (Li et al., 2014). Corroborating those preliminary findings, we have now determined that rare inherited PTVs in the entire CEACAM gene family are associated with breast cancer risk in both European and African Americans with respective p-values of $1.75 \times 10^{-13}$ and $1.87 \times 10^{-04}$. The p-value generated for African American breast cancer risk was likely influenced by the small sample size in TCGA.

We analyzed blood-derived exomes of European and African American breast cancer cases in TCGA to identify inherited PTVs in all human CEACAM genes, and detected sixteen and six rare PTVs in each ethnicity, respectively. Gene-based...
analyses determined that rare PTVs in CEACAM6, CEACAM7, and CEACAM8 are associated with European American breast cancer risk, and rare PTVs in CEACAM7 are associated with breast cancer risk in African Americans. CEACAM7, which was associated with breast cancer risk in both ethnicities, has no current link to breast cancer. However, down-regulation of CEACAM7 in hyperplastic polyps and early adenomas represent some of the earliest observable molecular events leading to colorectal tumors (Scholzel et al., 2000). Though CEACAM7 expression was thought to be restricted to the epithelial cells of the colon and pancreas, according to the Human Protein Atlas, granular cells of the breast have moderate CEACAM7 protein expression (Uhlen et al., 2015; Raj et al., 2021). How CEACAM7 plays a role in breast cancer is currently unknown, but the link could even be indirect and due to expression in non-breast tissue (Ferreira et al., 2019). CEACAM7 c.195C > A:p.(Y65X), which was detected in 10.8 and 4.5% of European and African American cases, respectively, was absent in all EVS controls. It severely truncates the 265 amino acid proteins and results in a loss of the cytoplasmic region, as well as a large portion of the extracellular region, including disruption of the Ig-like and Ig V-set domains. It is likely a loss-of-function mutation (Figure 3).

Rare PTVs in CEACAM6 and CEACAM8 appear to only be associated with European American breast cancer risk. Considering that CEACAM6/8 co-expression inhibits proliferation and invasiveness of breast cancer cells (Iwabuchi et al., 2019), having a rare PTV in one of those two genes may be sufficient to override their synergistic tumor-suppressing relationship. While a number of PTVs were detected in these genes, two splicing mutations, CEACAM6 c.*40 + 2T > G and CEACAM8 c.*40 + 2T > G, were individually determined to be associated with European American breast cancer, both of which affect non-coding exons in the 3' UTR. Both mutations affect the donor site immediately following exon 5 of their respective genes, which contains both coding and non-coding DNA. The mutated donor sites likely affect the downstream sequence of the mature mRNA product, either retaining (all or a part of) intron 5 or removing exon 6, the last non-coding exon, where many microRNA binding sites are located (Figure 4). Based on mirD8 rankings, the top five microRNAs that bind to the 3' UTRs of CEACAM6 and CEACAM8 have previous links to cancer (Table 4); thus, disrupted microRNA binding likely leads to aberrant CEACAM6 and CEACAM8 expression.

Two stop gain mutations in CEACAM4 (c.367C > T:p.R123X and c.424C > T:p.Q142X) were associated with African American breast cancer. These mutations were not detected in European American cases or controls, and are very rare in the general African American population. They were detected in significantly more African American breast cancer cases compared to ethnic-matched controls, suggesting their involvement in African American breast cancer risk. However, gene-based aggregation analyses did not support CEACAM4 as a breast cancer risk gene. Larger African American breast cancer cohorts will need to be studied to validate these findings. Interestingly, in a study of parous women with and without breast cancer, CEACAM4 has been reported to be up-regulated in normal breast compared to breast tumor samples (Balogh et al., 2007). Though race/ethnicity was not revealed in that study, the results suggest that CEACAM4 could be a breast cancer tumor suppressor.

It has long been reported that minimal genetic changes can have radical effects on the function of CEACAM genes (Naghbalihosaini and Stanners, 2004). Residues in CEACAM6 and CEACAM8 have been identified that are critical for CEACAM6 homodimerization as well as the formation of CEACAM6 and CEACAM8 heterodimers, which is important in preventing breast cancer cell proliferation (Kuroki et al., 2001; Iwabuchi et al., 2019). There have also been residues reported in CEACAM1 that are crucial for determining the risk of infection by receptor-binding pathogens (Villulas et al., 2007) and preventing the killing activity of NK cells (Markel et al., 2004). Furthermore, somatic missense mutations in colorectal cancers have been detected in CEACAM1 (Song et al., 2011) and CEACAM5 (Gu et al., 2020), the latter of which has been shown to increase proliferation by inhibiting TGFβ signaling and altering the intestinal microbiome. The microbiome has been reported as a new breast cancer risk factor (Fernandez et al., 2018; Eslami, 2020). In fact, differences have been reported in the microbiome of normal and cancerous breast tissue, as well as the gut microbiota of breast cancer cases versus controls (Fernandez et al., 2018). Disrupted CEACAM genes could be the underlying mechanism through altered TGFβ signaling, bacteria docking, and/or estrogen metabolism (Villulas et al., 2007; Tchoupa et al., 2014; Fernandez et al., 2018; Gu et al., 2020). This study reports the first association of inherited CEACAM mutations and breast cancer risk, and potentially implicates the whole gene family in genetic risk. Precisely how these mutations contribute to breast cancer needs to be determined, especially considering our current knowledge on the role that the CEACAM gene family plays in tumor development, progression, and metastasis.

DATA AVAILABILITY STATEMENT

The WGS data for the five whole genome sequenced CMT-affected Golden Retriever dogs can be obtained through the NCBI SRA repository through BioProject PRJNA745215. TCGA data is available through dbGAP.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Auburn University Institutional Review Board (IRB) for the Protection of Human Subjects in Research. The patients/participants provided their written informed consent to participate in this study. Ethical review and approval was not required for the animal study because this research did not require ORC – Animal Care & Use (IACUC) approval since only dog DNA was studied upon receipt from the CHIC repository.

AUTHOR CONTRIBUTIONS

AH and NM wrote the manuscript and performed variant and statistical analyses. AH and IM performed PCR for validation and
determining mutational frequency. AH performed bioinformatic processing. All authors read and approved the final manuscript.

FUNDING
This research was supported by the Department of Pathobiology in the Auburn University College of Veterinary Medicine and the Department of Drug Discovery and Development in the Auburn University Harrison School of Pharmacy. This work was partially funded by an Auburn University Research Initiative in Cancer (AURIC) Seed Grant for the canine WGS efforts. Graduate Student endeavors were supported by the AURIC Graduate Fellowship Program (to AH).

REFERENCES
Alshamrani, A. A. (2020). Roles of microRNAs in Ovarian Cancer Tumorigenesis: Two Decades Later, What Have We Learned? Front. Oncol. 10:1084.
American Cancer Society (2020). Cancer Facts & Figures 2020. New York, NY: American Cancer Society.
Amin Al Olama, A., Kote-Jarai, Z., Schumacher, F. R., Wiklund, F., Berndt, S. I., Benlloch, S., et al. (2013). A meta-analysis of genome-wide association studies to identify prostate cancer susceptibility loci associated with aggressive and non-aggressive disease. Hum. Mol. Genet. 22, 408–415.
Bai, Q. L., Hu, C. W., Wang, X. R., Shang, J. X., and Yin, G. F. (2017). MiR-616 promotes proliferation and inhibits apoptosis in glioma cells by suppressing expression of SOX7 via the Wnt signaling pathway. Eur. Rev. Med. Pharmacol. Sci. 21, 5630–5637.
Balogh, G. A., Russo, J., Mailo, D. A., Heulings, R., Russo, P. A., Morrison, P., et al. (2007). The breast of parous women without cancer has a different genomic profile compared to those with cancer. Int. J. Oncol. 31, 1165–1175.
Bamberger, A. M., Riethdorf, L., Nollau, P., Naumann, M., Erdmann, I., Gotze, J., et al. (1998). Dysregulated expression of CD66a (BGP, C-CAM), an adhesion molecule of the CEA family, in endometrial cancer. Am. J. Pathol. 152, 1401–1406.
Beauchemin, N., and Arzabadeh, A. (2013). Carcinoembryonic antigen-related cell adhesion molecules (CEACAMs) in cancer progression and metastasis. Cancer Metastasis Rev. 32, 643–671. doi: 10.1007/s10555-013-9444-6
Brackman, J. (2020). Large-Scale Cancer Study of Golden Retrievers Holds Hope for All Dogs. Available online at https://thebark.com/content/large-scale-cancer-study-golden-retrievers-holds-hope-all-dogs.(accessed date January 2016).
Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R. L., Torre, L. A., and Jemal, A. (2018). Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J. Clin. 68, 394–424. doi: 10.3322/caac.21492
Busch, C., Hanssen, T. A., and Wagener, C. (2002). Down-regulation of CEACAM1 in human prostate cancer: correlation with loss of cell polarity, increased proliferation rate, and Gleason grade 3 to 4 transition. Hum. Pathol. 33, 290–298. doi: 10.1053/hupa.2002.32218
Cardoso, F., Harbeck, N., Barrios, C. H., Bergh, J., Cortes, J., El Saghir, N., et al. (2017). Research needs in breast cancer. Ann. Oncol. 28, 208–217.
Carrtier, F., Inderse, E., Lesjean, S., Charpentier, J., Hooks, K. R., Ghousein, A., et al. (2017). New tumor suppressor microRNAs target glypican-3 in human liver cancer. Oncotarget 8, 41211–41226. doi: 10.18632/oncotarget.17162
Chen, C., Xue, S., Zhang, J., Chen, W., Gong, D., Zheng, J., et al. (2017). DNA-methylation-mediated repression of miR-766-3p promotes cell proliferation via targeting SFP2 expression in renal cell carcinoma. Int. J. Cancer 141, 1867–1878. doi: 10.1002/ijc.310853
Chen, F., Zhou, H., Wu, C., and Yan, H. (2018). Identification of miRNA profiling in prediction of tumor recurrence and progress and bioinformatics analysis for patients with primary esophageal cancer: Study based on TCGA database. Pathol. Res. Pract. 214, 2081–2086. doi: 10.1016/j.prp.2018.10.009
Chen, Y., and Wang, X. (2020). miRDB: an online database for prediction of functional microRNA targets. Nucleic Acids Res. 48, D127–D131.
Chen, Z., Zhu, J., Zhu, Y., and Wang, J. (2018). MicroRNA-616 promotes the progression of ovarian cancer by targeting TIMP2. Oncol. Rep. 39, 2960–2968.
Dobson, J. M. (2013). Breed-predispositions to cancer in pedigree dogs. ISRN Vet. Sci. 2013:941275.
Deli, S. Z., Majidzadeh, A. K., Halvazi, S., Babapirali, F., and Esmaeili, R. (2020). Microbiome and Breast Cancer: New Role for an Ancient Population. Front. Oncol. 10:120.
Estari, M. A., Esmaeili, R., Zare, A. A., Farahmand, I., Fazlizad, H., Zeki, A., et al. (2017). High expression of CEACAM19, a new member of carcinoembryonic antigen gene family, in patients with breast cancer. Clin. Exp. Med. 17, 547–553. doi: 10.1007/s10238-016-0442-1
Ettlin, J., Clementi, E., Amini, P., Malbon, A., and Markkanen, E. (2017). Analysis of Gene Expression Signatures in Cancer-Associated Stroma from Canine Mammary Tumours Reveals Molecular Homology to Human Breast Carcinomas. Int. J. Mol. Sci. 18:5.
Exome Variant Server (2019). NHLBI GO Exome Sequencing Project (ESP). Available online at: https://evs.gs.washington.edu/ESVS/ (accessed November 2, 2020).
Fernandez, M. F., Reina-Perez, L., Astorga, J. M., Rodriguez-Carrillo, A., Plaza-Diaz, J., and Fontana, L. (2018). Breast Cancer and Its Relationship with the Microbiota. Int. J. Environ. Res. Public Health 15:8.
Ferreira, M. A., Gamazon, E. R., Al-Ejeh, F., Aittomaki, K. I, Andrulis, L., Anton-Culver, H., et al. (2019). Genome-wide association and transcriptome studies identify target genes and risk loci for breast cancer. Nat. Commun. 10:1741.
Fishar, R. A. (1925). Statistical methods for research workers. Edinburgh: Oliver and Boyd.
Fournes, B., Sadekova, S., Turbide, C., Letourneau, S., and Beauchemin, N. (2001). The CEACAM1-L Ser503 residue is crucial for inhibition of colon cancer cell tumorigenicity. Oncogene 20, 219–230. doi: 10.1038/sj.onc.120458
Fyfe, J. C., Hemker, S. L., Frampton, A., Raj, K. N., Nagy, P. L., Gibbon, K. J., et al. (2018). Inherited selective cobalamin malabsorption in Komondor dogs associated with a CUBN splice site variant. BMC Vet. Res. 14:418.
Gao, P., Xia, J. H., Sipeky, C., Dong, X. M., Zhang, Q., Yang, Y., et al. (2018). Biological and Clinical Implications of the 19q13 Aggressive Prostate Cancer Susceptibility Locus. Cell 174, 576–589.e518.
Garden, O. A., Volk, S. W., Mason, N. J., and Perry, J. A. (2018). Companion animals in comparative oncology: One Medicine in action. Vet. J. 240, 6–13. doi: 10.1016/j.tvjl.2018.08.008
Gaur, S., Shively, J. E., Yen, Y., and Gaur, R. K. (2008). Altered splicing of CEACAM1 in breast cancer: identification of regulatory sequences that control splicing of CEACAM1 into long or short cytoplasmic domain isoforms. Mol. Cancer 7:46. doi: 10.1186/1476-4598-7-46
Ge, S., Sun, C., Hu, Q., Guo, Y., Xia, G., Mi, Y., et al. (2020). Differential expression profiles of circRNAs in human prostate cancer based on chip and bioinformatic analysis. Int. J. Clin. Exp. Pathol. 13, 1045–1052.

ACKNOWLEDGMENTS
We would like to acknowledge the Orthopedic Foundation for Animals’ CHIC DNA Repository, which provided CMT-affected dog DNA samples. We would also like to thank the Office of Information Technology at Auburn University Hopper High-Performance Computing Cluster for computer time and technical support.

SUPPLEMENTARY MATERIAL
The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2021.702889/full#supplementary-material
Kim, K. K., Seung, B. J., Kim, D., Park, H. M., Lee, S., Song, D. W., et al. (2019). Whole-exome and whole-transcriptome sequencing of canine mammary gland tumors. Sc. Data 6:147.

Kolicheski, A. L., Johnson, G. S., Mhlanga-Mutangadura, T., Taylor, J. F., Schnabel, R. D., Kinosita, H., et al. (2017). A homologous PIK3CA missense mutation in Soft-Coated Wheaten Terriers with a canine paroxysmal dyskinesia. Neurogenetics 18, 39–47. doi: 10.1007/s10048-016-0502-4

Kuespert, K., Pilis, S., and Hauck, C. R. (2006). CEACAMs: their role in physiology and pathophysiology. Curr. Opin. Cell Biol. 18, 565–571. doi: 10.1016/j.ceb.2006.08.008

Kumar, M., Gouw, M., Michael, S., Samano-Sanchez, H., Pancsa, R., Glavina, J., et al. (2020). ELM-the eukaryotic linear motif resource in 2020. Nucleic Acids Res. 48, D296–D306.

Kuroki, M., Abe, H., Imakirei, T., Liao, S., Uchida, H., Yamauchi, Y., et al. (2001). Identification and comparison of residues critical for cell-adhesion activities of two neuropilin CD66 antigens, CEACAM6 and CEACAM8. J. Leukoc Biol. 70, 543–550.

Lee, K. H., Hwang, H. J., Noh, H. J., Shin, T. J., and Cho, J. Y. (2019). Somatic Mutation of PIK3CA (H1047R) Is a Common Driver Mutation Hotspot in Canine Mammary Tumors as Well as Human Breast Cancers. Cancers 11:12.

Lee, K. H., Park, H. M., Son, K. H., Shin, T. J., and Cho, J. Y. (2018). Transcriptome Signatures of Canine Mammary Gland Tumors and Its Comparison to Human Breast Cancers. Cancers 10:9.

Li, H., Yang, B., Xing, K., Yuan, N., Wang, B., Chen, Z., et al. (2014). A preliminary study of the relationship between breast cancer metastasis and loss of heterozygosity by using exome sequencing. Sci. Rep. 4:5460.

Li, W., Li, Y., Ma, W., Zhou, J., Sun, Z., and Yan, X. (2020). Long noncoding RNA AC114812.8 promotes the progression of bladder cancer through miR-371b-5p/FUT4 axis. Biomed. Pharmacother. 121:109605. doi: 10.1016/j.biopha.2019.109605

Liu, D., Xiong, H., Ellis, A. E., Northrup, N. C., Rodriguez, C. O., O'Regan, R. M., et al. (2014). Molecular homology and difference between spontaneous canine mammary cancer and human breast cancer. Cancer Res. 74, 5045–5056. doi: 10.1158/0008-5472.can-14-0392

Liu, F., Cai, Y., Rong, X., Chen, J., Zheng, D., Chen, L., et al. (2017). MiR-661 promotes tumor invasion and metastasis by directly inhibiting RBP1 in non small cell lung cancer. Mol. Cancer 16:122.

Liu, J., Muturi, H. T., Khuder, S. A., Helal, R. A., Ghadieh, H. E., Ramakrishnan, S. K., et al. (2020). Loss of Ceacam1 promotes prostate cancer progression in Pten haploinsufficient male mice. Metabolism 107:154215. doi: 10.1016/j.metabol.2020.154215

Liu, S., Lin, Z., Zheng, Z., Rao, W., Lin, Y., Chen, H., et al. (2020). Serum exosomal microRNA-766-3p expression is associated with poor prognosis of esophageal squamous cell carcinoma. Cancer Sci. 111, 3881–3892. doi: 10.1111/cas.14530

Luo, X., Zhang, X., Peng, J., Chen, Y., Zhao, W., Jiang, X., et al. (2020). miR-371b-5p promotes cell proliferation, migration and invasion in non-small cell lung cancer via SCAL1. Biosci. Rep. 40:11.

Madeira, F., Park, Y. M., Lee, J., Buso, N., Guri, T., Madhusoodanan, C. N. et al. (2019). The EMBL-EBI search and sequence analysis tools APIs in 2019. Nucleic Acids Res. 47, W636–W641.

Maraqa, L., Cummings, M., Peter, M. B., Shaaban, A. M., Horgan, K., Hanby, A. M., et al. (2008). Carcinoembryonic antigen cell adhesion molecule 6 predicts breast cancer recurrence following adjuvant tamoxifen. Clin. Cancer Res. 14, 405–411. doi: 10.1158/1078-0432.ccr-07-1363

Markel, G., Gruda, R., Achdout, H., Katz, G., Nechama, M., Blumberg, R. S., et al. (2004). The critical role of residues 43R and 44Q of carcinoembryonic antigen-related cell adhesion molecule 6 and 8 inhibits proliferation and invasiveness of breast carcinoma cells. Clin. Exp. Metastasis 34, 423–432. doi: 10.1007/s10585-019-09981-2

Jiang, X., Jiang, M., Xu, M., Xu, J., and Li, Y. (2019). Identification of diagnostic utility and molecular mechanisms of circulating miR-551b-5p in gastric cancer. Pathol. Res. Pract. 215, 900–904. doi: 10.1016/j.prp.2019.01.035

Kammerer, R., and Zimmermann, W. (2010). Coevolution of activating and inhibitory receptors within mammalian carcinoembryonic antigen families. BMC Biol. 8:12.

Kammerer, R., Popp, T., Hartle, S., Singer, B. B., and Zimmermann, W. (2007). Species-specific evolution of immune receptor tyrosine based activation motif-containing CEACAM1-related immune receptors in the dog. BMC Evol Biol 7:196. doi: 10.1186/1471-2148-7-196

Kammerer, R., Rüttiger, L., Riesenberg, R., Schaubel, C., Krupar, R., Kamp, A., et al. (2012). Loss of mammal-specific tectorial membrane component carcinoembryonic antigen cell adhesion molecule 16 (CEACAM16) leads to hearing impairment at low and high frequencies. J. Biol. Chem. 287, 21584–21598. doi: 10.1074/jbc.m112.320481

Keast, M. S., Burton, J. H., Dank, G., Bannasch, D. L., and Rebhun, R. B. (2018). Association of cancer-related mortality, age and gonadectomy in golden retriever dogs at a veterinary academic center (1989-2016). PLoS One 13:e0192578. doi: 10.1371/journal.pone.0192578
Yasui, T., Yanagida, T., Ito, S., Konakade, Y., Takeshita, D., Naganawa, T., et al. (2017). Unveiling massive numbers of cancer-related urinary-microRNA candidates via nanowires. Sci. Adv. 3:e1701133. doi: 10.1126/sciadv.1701133

Yin, J., Rockenbauer, E., Hedayati, M., Jacobsen, N. R., Vogel, U., Grossman, L., et al. (2002). Multiple single nucleotide polymorphisms on human chromosome 19q13.2-3 associate with risk of Basal cell carcinoma. Cancer Epidemiol Biomarkers Prev 11, 1449–1453.

Yokoi, A., Matsuzaki, J., Yamamoto, Y., Tate, K., Yoneoka, Y., Shimizu, H., et al. (2019). Serum microRNA profile enables preoperative diagnosis of uterine leiomyosarcoma. Cancer Sci. 110, 3718–3726. doi: 10.1111/cas.14215

You, Y., Que, K., Zhou, Y., Zhang, Z., Zhao, X., Gong, J., et al. (2018). MicroRNA-766-3p Inhibits Tumour Progression by Targeting Wnt3a in Hepatocellular Carcinoma. Mol. Cells 41, 830–841.

Zerhino, D. R., Achuthan, P., Akanni, W., Amode, M. R., Barrell, D., Bhai, J., et al. (2018). Ensembl 2018. Nucleic Acids Res. 46, D754–D761.

Zhang, C., Xue, Q., Xu, Z., and Lu, C. (2018). MiR-5702 suppresses proliferation and invasion in non-small-cell lung cancer cells via posttranscriptional suppression of ZEB1. J. Biochem. Mol. Toxicol. 2018:e22163. doi: 10.1002/jbt.22163

Zhang, S., Chen, H., Liu, W., Fang, L., Qian, Z., Kong, R., et al. (2020). miR-766-3p Targeting BCL9L Suppressed Tumorigenesis, Epithelial-Mesenchymal Transition, and Metastasis Through the beta-Catenin Signaling Pathway in Osteosarcoma Cells. Front. Cell Dev. Biol. 8:394135.

Zheng, J., Miller, K. K., Yang, T., Hildebrand, M. S., Shearer, A. E., DeLuca, A. P., et al. (2011). Carcinoembryonic antigen-related cell adhesion molecule 16 interacts with alpha-tectorin and is mutated in autosomal dominant hearing loss (DFNA4). Proc. Natl. Acad. Sci. U S A 108, 4218–4223. doi: 10.1073/pnas.1005842108

Zhu, L. M., and Li, N. (2020). Downregulation of long noncoding RNA TUSC7 promoted cell growth, invasion and migration through sponging with miR-616-3p/GSK3beta pathway in ovarian cancer. Eur. Rev. Med. Pharmacol. Sci. 24, 7253–7265.

Zhu, T., Yuan, J., Wang, Y., Gong, C., Xie, Y., and Li, H. (2015). MiR-661 contributed to cell proliferation of human ovarian cancer cells by repressing INPP5J expression. Biomed. Pharmacother. 75, 123–128. doi: 10.1016/j.biopha.2015.07.023

Zisi, Z., Adamopoulos, P. G., Kontos, C. K., and Scorilas, A. (2020). Identification and expression analysis of novel splice variants of the human carcinoembryonic antigen-related cell adhesion molecule 19 (CEACAM19) gene using a high-throughput sequencing approach. Genomics 112, 4268–4276. doi: 10.1016/j.ygeno.2020.06.043

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Huskey, McNeely and Merner. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.