

IDENTIFICAÇÃO DE BIOMARCADORES EM SORO SANGUÍNEO PARA DIAGNÓSTICO DE NEOPLASIA EM CÃES POR ESPECTROSCOPIA RAMAN

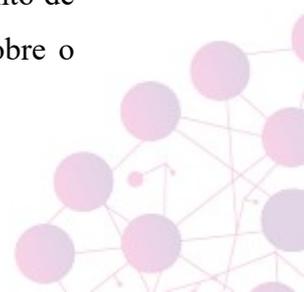
Msc. Daniela Franco Lopes¹, Dr. Landulfo Silveira Jr² (orientador)

Resumo:

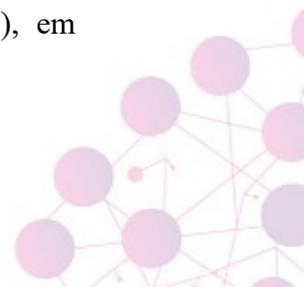
A detecção de biomarcadores de câncer no soro pela da E.R a 830 nm foi aplicada com o objetivo de identificar biomarcadores em neoplasias no soro de 61 cães e a discriminação do grupo a que pertencente por análise estatística multivariada. Os 354 espectros foram submetidos à PCA para identificar as diferenças espectrais e correlacionar com as diferenças de composição no soro dos grupos Saudável e Câncer, e classificação dos espectros utilizou a PLS. A PCA revelou picos em 1132, 1342, 1368 e 1453 cm^{-1} , "redshift" em 621 (617/624), 1003 (1000/1006) e 1032 (1028/1034) cm^{-1} , elevados no grupo câncer, e picos a 451 cm^{-1} e 1441 cm^{-1} elevados no grupo saudável. O PLS-DA (por paciente) classificou os espectros do soro com 78,0% de precisão, 76,0% de sensibilidade e 81,0% de especificidade. Os resultados sugerem a presença de biomarcadores relacionados com alterações em neoplasia no soro de cães que podem ser detectados por espectroscopia Raman.

Introdução:

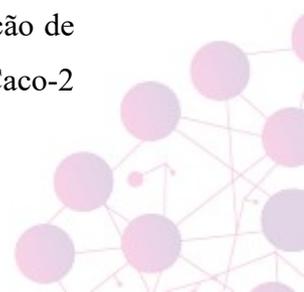
Em um estudo de oncologia comparativa realizado nos EUA, verificou-se que cerca de 1,66 milhões de humanos e mais de 4,2 milhões de cães são diagnosticados com câncer anualmente (SCHIFFMANN; BREEN, 2015). A ocorrência de tumores em cães e gatos os torna modelos adequados para o estudo da gênese tumoral em humanos, pois possuem uma particularidade para desenvolverem tumores espontâneos e a semelhança genômica humana em linhagens germinativas (SCHIFFMANN; BREEN, 2015). Estas predisposições ao câncer em humanos foram encontradas no DNA de cães, por exemplo, os genes BRCA-1 e BRCA-2 e TP-53 (SCHIFFMANN; BREEN, 2015), aumentando o interesse no desenvolvimento de estudo translacional em oncologia comparativa que pode fornecer insights sobre o



risco e a prevenção do câncer em humanos (SCHIFFMANN; BREEN, 2015; BODDY, et al., 2020). No entanto, o conhecimento da prevalência e mortalidade por câncer em animais é deficitário e depende potencialmente dos estudos em humanos (BODDY, et al., 2020). Provavelmente por estarem tão próximos dos humanos, os animais também sofrem exposição aos mesmos fatores de risco classificados pela *International Agency for Research on Cancer* (IARC) como radiações e intoxicações químicas por metais pesados, agentes biológicos e hormônios que poderiam ser evitáveis ao impedir estas exposições (SZENTI-GYORGYI, 1977; GROSSE et al., 2019). Apesar da evolução científica, não é possível obter a excelência em diagnóstico precoce do câncer para animais, pois muitos tipos de câncer são assintomáticos. O diagnóstico é baseado na história clínica e na anamnese, associados aos exames laboratoriais e complementares de imagem como ultrassom, raios-X, tomografia computadorizada e ressonância magnética (SILVA, 2019; WISNER; POLLARD, 2004; HAGE; MASAO, 2009), que auxiliam no prognóstico e monitoramento do paciente. A confirmação da doença ocorre após a biópsia histopatológica, considerada padrão ouro (BRAZ et al., 2016). A necessidade em realizar um diagnóstico precoce e rápido é uma das desvantagens no padrão ouro devido à lentidão no tempo de processamento da análise, a necessidade de exames prévios e anestesia geral e que podem onerar o diagnóstico e atrasar o início do tratamento, resultando no avanço da doença oncológica, sem permitir a monitoração em tempo real. Técnicas ópticas como a espectroscopia Raman têm sido propostas como um método de diagnóstico minimamente invasivo e com rapidez para verificar as alterações patológicas em tecidos (SILVEIRA et al., 2012; HANLON et al., 2000) e fluidos biológicos (HANLON et al., 2000; DUARTE et al., 2012; BORGES et al., 2015) de forma que o espectro Raman identifique os componentes bioquímicos que diferenciam um paciente com câncer de um indivíduo saudável, e com isso detectar a doença neoplásica e sugerir um diagnóstico (SILVEIRA et al., 2012; HANLON et al., 2000; BIRTOIU et al., 2016; GONZÁLEZ-SOLÍS et al., 2014; MELITTO et al., 2022). A espectroscopia Raman pode caracterizar e diferenciar tecidos normais de tecidos benignos, malignos e pré-malignos em cervix (GONZÁLEZ-SOLÍS et al., 2014), em mama (CHOWDARY et al., 2016), em



cérebro (BANERJEE; ZHANG, 2007), em trato gastrointestinal (BOHORFOUSH, 1996) em humanos e tecidos tumoral mamário em humanos e animais (BIRTOIU et al., 2016), além de auxiliar na detecção do vírus da COVID-19 (GOULART et al., 2022) e imunoglobulinas presentes na toxoplasmose em gatos (DUARTE et al., 2012) utilizando uma quantidade mínima de amostra de sangue. A espectroscopia Raman se baseia no espalhamento inelástico da luz que incide nas moléculas (HANLON et al., 2000), possibilitando a determinação da composição bioquímica nos aspectos quantitativos (concentração) bem como nos qualitativos (discriminação e classificação) (SILVEIRA et al., 2012; HANLON et al., 2000). as vantagens da espectroscopia Raman em comparações a outras técnicas de diagnóstico óptico como a espectroscopia de absorção incluem a obtenção de informações da energia vibracional das moléculas (HANLON et al., 2000) em tempo real, sem necessidade de preparação específica da amostra (MOREIRA et al., 2008). esta técnica pode ser empregada no diagnóstico *in situ* e *ex situ* de alterações neoplásicas e de forma complementar as informações obtidas pela histopatologia (BRAZ et al., 2016) e imagens (SILVA, 2019; WISNER; POLLARD, 2004; HAGE; MASAO, 2009), podendo auxiliar na monitoração das doenças e monitoramento da quimioterapia em leucemias (GONZÁLEZ-SOLÍS et al., 2014). As primeiras investigações utilizando a espectroscopia Raman para pesquisa do sangue humano iniciaram em 1970, onde os espectros Raman (excitação de 514,5 nm) da membrana eritrocitária íntegra foram comparadas com os espectros de membranas eritrocitárias sem as proteínas periféricas (GOHEEN, et al., 1977). Na oncologia, González-Solís et al., (2014) utilizaram a espectroscopia Raman e a análise de componente principal (PCA) para diferenciação e classificação de soro de indivíduos leucêmicos dos não leucêmicos. Martínez-Espinoza et al., (2008) utilizaram a espectroscopia Raman com a PCA para diferenciar amostras de soro sanguíneo de pacientes com leucemia dos normais e discriminar a quais grupos pertencem. González-Solís et al., (2014) utilizaram a espectroscopia Raman para analisar a composição bioquímica de amostras de soro e distinguir entre indivíduos normais e com câncer cervical usando a PCA para a discriminação dos grupos com alta sensibilidade e especificidade. A aplicação da espectroscopia Raman *in vitro* possibilitou a investigação e identificação da ação de medicamentos homeopáticos na viabilidade mitocondrial de células Caco-2



sugerindo ação citotóxica promovida pelo *Viscum album* (LOPES et al., 2009). Na área veterinária, MUNTEANU et al., (2017) utilizaram a espectroscopia Raman amplificada na superfície (SERS) para o diagnóstico de tecidos tumorais de cães e comparando os resultados obtidos com a biópsia histopatológica, demonstrando que a técnica empregada foi capaz de distinguir tecidos malignos de tecidos normais. Duarte et al., (2010) demonstraram um modelo de diagnóstico óptico baseado na detecção de imunoglobulinas em soro sanguíneo de felinos usando a técnica espectroscopia Raman associada ao modelo de regressão por mínimos quadrados parciais (PLS), obtendo alta especificidade e sensibilidade para o diagnóstico de toxoplasmose comparativamente ao padrão ouro fornecido pelo método ELISA, com potencial para se tornar uma ferramenta para o uso na Medicina Veterinária. Este estudo preliminar teve como objetivo utilizar a espectroscopia Raman para identificar as diferenças espectrais em amostras de soro sanguíneo de cães entre pacientes de grupos Saudável e Câncer e aplicar a técnica PCA em análise exploratória para identificar as prováveis diferenças na composição bioquímica de proteínas, aminoácidos e lipídeos entre esses grupos, que estão relacionados a alterações presentes em doenças neoplásicas. Propôs-se também a classificação de cada um dos espectros nos grupos Saudáveis e Câncer por análise discriminante via PLS utilizando a validação cruzada “leave one out”, onde a PLS utiliza a diferença na composição bioquímica dos grupos como variável de classificação ou diferenciação entre os grupos.

Palavras-Chave: biomarcadores tumorais; soro sanguíneo; análise discriminante; análise de componentes principais; regressão por mínimos quadrados parciais; espectroscopia Raman

Métodos:

Amostras de soro sanguíneo de cães domésticos: A O estudo foi aprovado pelo Comitê de Ética e Pesquisa da Universidade Anhembí Morumbi (Protocolo No. 05001/2021) com Consentimento Livre e Esclarecido assinado pelos tutores dos cães. Amostras de sangue cefálico foram coletadas de cada paciente, acondicionadas em tubo seco sem anticoagulante (modelo BD Vacutainer®, Interlab Distribuidora de Produtos Científicos Ltda., São Paulo, SP, Brasil) e enviadas ao Animaltec – Centro



de Diagnósticos Veterinários (Campinas, SP, Brasil), utilizando centrífuga (modelo Elektra CE132, Labor Line Ind. Com. e Serviços Ltda. – ME, Osasco, SP, Brasil). Os tubos com as amostras continham 3 mL de sangue e centrifugados a 1500 rpm por 15 minutos para separação do soro sanguíneo e imediatamente armazenados sob refrigeração à 8 °C por 48 h ou congelados à –20 °C por 7 dias para minimizar alterações na constituição bioquímica do soro até a realização do experimento para coleta dos espectros. A Tabela 1 apresenta o número de animais inscritos no estudo por tipos de tumores e incidência. As amostras de soro dos 61 pacientes foram separadas em dois grupos: 31 amostras de animais clinicamente saudáveis (grupo Saudável), e 30 animais diagnosticados com câncer com 41 amostras (grupo Câncer), onde o diagnóstico foi feito através da biópsia histopatológica e dos exames de imagem. Em alguns pacientes do grupo Câncer foi realizado mais de uma coleta de sangue e/ou tinham mais de um tipo de tumor. No grupo Câncer prevaleceu fêmeas acima de nove anos, com raça, e os tumores de maior ocorrência foram tumores esplênicos (nove amostras), tumores de mama (cinco amostras) e lipomas (quatro amostras), enquanto no grupo Saudável prevaleceram fêmeas na faixa etária de dois a oito anos (Tabela 2). Os pacientes oncológicos não receberam terapia anterior à coleta. Foram excluídos da análise de dados sete pacientes com sintomas clínicos sugestivos de tumor e cujas necropsias indicaram óbito em decorrência de neoplasias, porém não possuíam resultado confirmatório por biópsia ou imagem (Tabela 1).

Tabela 1. Number of cancer patients enrolled in the study by type of tumors, number of samples of each tumor type and method of diagnosis.

Type of tumors	Number of patients	Number of samples	Report
Amelanotic melanoma	1	3	Histopathology
Breast carcinoma	2	3	Histopathology
Epithelioma	1	1	Histopathology
Hepatic adenoma	1	1	Histopathology



Histiocytoma	1	1	Histopathology
Lipoma	1	2	Histopathology
Mastocytoma	1	1	Histopathology
Sarcoma	3	3	Histopathology
Squamous cell carcinoma	2	3	Histopathology
Suggestive of spleen hemangiosarcoma	5	9	US
Suggestive of brain tumor	3	4	X-Ray, US, CT
Suggestive of renal cyst	2	2	US
Suggestive of pancreas tumor	3	3	US
Suggestive myeloma	3	3	X-Ray, CT
Suggestive pheochromocytoma	1	1	US
Suggestive of bladder tumor	1	1	US
Suggestive of neoplasia*	7	7	Clinical examination
Total	30	41	

Espectroscopia Raman. Para a avaliação espectroscópica, o soro dos animais foi descongelado passivamente 10 minutos antes da avaliação e uma alíquota de 80 μL de soro sem preparo foi pipetada em um porta amostras de alumínio com furos utilizando uma micropipeta (modelo P200, Bio-Rad Laboratories Inc., Hercules, CA, EUA). Os espectros foram coletados em quintuplicata para fins estatísticos. Os espectros foram adquiridos no espectrômetro Raman dispersivo (modelo Dimension-P1, Lambda Solutions Inc., Waltham, MA, USA) com excitação de 830 nm, potência do laser ajustada para 350 mW, e tempo de exposição de 3 s com 10 acumulações (30 s de tempo total). O espectrômetro possui resolução de 2 cm^{-1} na faixa espectral útil de 400 a 1800 cm^{-1} . A amostra foi iluminada e o sinal Raman foi coletado por

meio de uma sonda Raman acoplada ao espectrômetro. Após a coleta, os raios cósmicos foram removidos manualmente e os espectros foram submetidos a um pré-processamento com a finalidade de corrigir a linha de base (devido principalmente à autofluorescência), sendo empregado um polinômio de ordem 7 que foi ajustado e subtraído de cada espectro. Os espectros foram normalizando pela área sob a curva (1-norm) (LASCH, 2012). O pré-processamento foi realizado utilizando uma rotina desenvolvida no software MATLAB (versão 2007a, The MathWorks Inc., Natick, MA, USA). Os espectros médios de cada grupo foram calculados e plotados para comparação visual e posterior identificação das diferenças nas intensidades dos picos Raman que caracterizem as diferenças de constituição entre os grupos Saudável e Câncer. Os picos mais intensos em ambos os espectros foram marcados e suas posições anotadas, a fim de determinar a composição do soro sanguíneo baseado nas posições dos picos Raman da literatura; o espectro da diferença também foi calculado e os picos da diferença também foram anotados.

Análise Exploratória por PCA. É uma ferramenta de análise multivariada que pode ser utilizada para análise exploratória dos dados espectrais. A PCA realiza a compactação dos dados em amostras grandes e que contenham múltiplas variáveis concomitantemente com a estimativa da variância do conjunto de dados (JOLLIFFE; CADIMA, 2016). Com base nessa variância, a PCA determina novos eixos (os componentes principais) que maximizam essa variância e apresentam ortogonalidade entre si. O primeiro componente principal corresponde ao eixo de maior variância espectral; o segundo componente principal aponta para a direção da segunda maior variância e não está correlacionado com o primeiro componente (ortogonal) (JOLLIFFE; CADIMA, 2016), e os componentes subsequentes estimam as variâncias restantes. A intensidade de cada um desses eixos nos dados originais também é registrada. Portanto, a PCA gera duas novas variáveis: os escores dos componentes principais (Scores), que são as direções da variância, e os respectivos carregamentos (PCs), que indicam a intensidade de cada Score dos dados originais. Os Scores são similares a espectros Raman e podem ser empregados para indicar a constituição bioquímica das amostras de soro e suas diferenças em função do grupo (Saudável ou Câncer), enquanto os PCs apontam onde (em qual grupo) estão essas



diferenças (JOLLIFFE; CADIMA, 2016, SILVA et al., 2018). A finalidade da análise exploratória via PCA foi identificar nos Scores os picos de compostos bioquímicos que estão presentes no soro e determinar se existem diferenças estatisticamente significativas nas intensidades dos PCs entre os grupos Saudável e Câncer, sugerindo uma classificação através das diferenças na constituição bioquímica entre os grupos e a presença de biomarcadores de neoplasia no grupo Câncer. As intensidades dos PCs foram avaliadas estatisticamente para indicar a significância estatística da hipótese de igualdade na média das intensidades dos PCs entre os grupos Saudável e Câncer (hipótese nula), com a aceitação ($p\text{-valor} > 0.05$) ou rejeição ($p\text{-valor} < 0.05$) da hipótese nula (que representa a igualdade na média das intensidades dos picos dos referidos grupos), onde as diferenças foram consideradas significativas quando o $p\text{-valor} < 0.05$ (VARGAS-OBIETA et al., 2016). Foi aplicado o teste de normalidade da distribuição das intensidades dos PCs utilizando o teste de Kolmogorov Smirnov. Nos PCs com distribuição gaussiana, foi aplicado o teste t de Student e quando a distribuição foi não-gaussiana foi aplicado o teste de Mann-Whitney. A PCA foi calculada usando a rotina “princomp.m” no software Matlab (versão 2007a, The MathWorks Inc., Natick, MA, EUA) (NUNES et al., 2012) e a análise estatística dos PCs foi calculada utilizando o software InStat (versão 3.00, GraphPad, San Diego, CA, USA, www.graphpad.com).

Análise discriminante por regressão PLS. A regressão PLS é também uma técnica estatística multivariada que pode ser empregada em análise discriminante (PLS-DA) para a classificação de amostras com muitas variáveis envolvidas (NGUYEN; ROCKE, 2002). A finalidade desta análise é discriminar amostras em grupos com base em um conjunto de dados de treinamento (NGUYEN; ROCKE, 2002; BARKER; RAYENS, 2007), onde as informações de variação estão presentes nos espectros Raman e que se correlacionam com os grupos. A PLS-DA foi utilizada em discriminação de leucemias (DA SILVA et al., 2018; LIMA et al., 2022), carcinoma basal celular, melanoma (SILVEIRA et al., 2012) e no diagnóstico de subtipos moleculares de câncer de mama em humanos (MELITTO et al., 2022). Através de um modelo de classificação dos espectros nos grupos Saudáveis e Câncer, onde as categorias preditas representam os grupos que se pretende discriminar: Saudável, grupo 1, e Câncer, grupo 2, e os preditores são os espectros Raman, as variáveis do

conjunto de dados espectrais são maximizadas pela covariância entre o predito e o preditor, e o modelo de classificação utiliza essas variáveis para prever a que grupo pertence uma amostra que foi aleatoriamente retirada do modelo (validação cruzada “leave-one-out”) (NUNES et al., 2012). O número de variáveis latentes (LVs) usado no modelo PLS-DA é o que apresenta a maior acurácia de discriminação entre os grupos Saudável e Câncer. Com os resultados da classificação apresentados em uma tabela de contingência, foram calculadas a acurácia, a sensibilidade e a especificidade de discriminação analisando os resultados da classificação na forma “spectrum wise”, ou seja, cada espectro representou um diagnóstico em um dos grupos, e na forma “patient wise”, onde o resultado da classificação de um paciente é o resultado da classificação da maioria simples dos espectros daquele paciente em um dos grupos. O modelo PLS-DA foi desenvolvido utilizando o software Chemoface (versão 1.65, UFLA, MG, Brazil, www.ufla.br/chemoface/) (NUNES et al., 2012).

Resultados e Discussões

Espectro do soro sanguíneo: A composição fisiológica do soro sanguíneo de mamíferos é complexa e similar entre as espécies, contendo água, sais inorgânicos, e biomoléculas como proteínas, aminoácidos, lipoproteínas, vitaminas, hormônios e glicose (LIMA et al., 2022). Estas biomoléculas são sensíveis às mudanças homeostáticas pré-cancerígenas e cancerígenas, por isso é de interesse aplicar a espectroscopia Raman para o diagnóstico e monitoramento do câncer (GONZÁLEZ-SOLÍS et al., 2014), uma vez que esta doença eleva as concentrações de proteínas séricas (albumina e globulinas) (PARACHALIL et al., 2019; DURAN et al., 2014; DO.N.S. LIMA et al., 2021). Os espectros Raman médios dos soros dos cães (Figura 1) possuem picos em 455, 506, 621, 642, 757, 853, 939, 1003, 1032, 1130, 1209, 1272, 1315, 1344, 1360, 1453 e 1656 cm^{-1} , atribuídos a proteínas séricas como a albumina e globulina (MARTINEZ-ESPINOSA et al., 2008; PARACHALIL et al., 2019; LIN; KONIG, 1976). Os picos com maior intensidade (1003, 1453 e 1656 cm^{-1}) estão mais intensos no grupo com Câncer, enquanto, o pico em 455 cm^{-1} está mais intenso no grupo Saudável. Estes picos constituem a maior parte do soro e podem ser usadas como biomarcadores (MARTINEZ-ESPINOSA et al., 2008). O espectro da



diferença entre os dois grupos (Câncer – Saudável, Figura 1) os picos positivos em 1132, 1342, 1368 e 1453 cm^{-1} podem ser atribuídos a proteínas e/ou aminoácidos e sugerem concentrações maiores no grupo Câncer, os picos negativos em 451 cm^{-1} , atribuído a triptofano (DE GELDER et al., 2007), e em 1441 cm^{-1} , atribuído a lipídeos[câncer (GONZÁLEZ-SOLÍS et al., 2014; KRAFFT et al., 2005; VARGAS-OBIETA et al, 2016), sugerem concentrações maiores no grupo Saudável. Os picos negativos/positivos em 617/624, 1000/1006 e 1028/1034 cm^{-1} sugerem “red shift” ou deslocamento para a direita (frequências maiores) dos picos em 621, 1003 e 1032 cm^{-1} (fenilalanina/tirosina) (DE GELDER et al., 2007) e esse deslocamento está presente nos espectros do grupo Câncer. Uma análise mais detalhada destes picos será apresentada na análise exploratória a seguir no próximo item.

População (n = 61)	M	F	Raça	SRD	< 1 ano	2 8 ano	> 9 anos
Healthy (n = 31)	15	16	22	9	4	17	10
Cancer (n = 30)	9	21	23	7	0	6	24

Tabela 1 – Número de pacientes incluídos no estudo por sexo, raça e idade

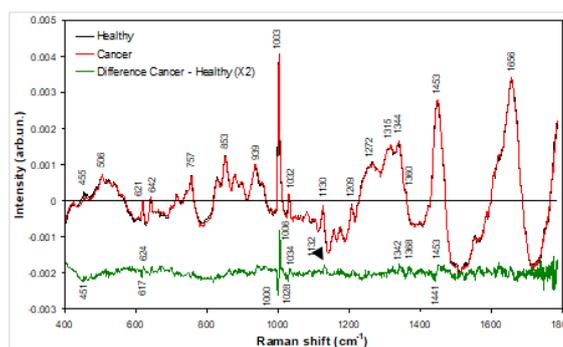


Fig.1- Espectros Raman médios do soro sanguíneo de indivíduos dos grupos Saudável e Câncer e o espectro da diferença (Câncer - Saudável). A linha verde representa os picos que são mais intensos nos espectros do grupo Câncer do que nos espectros do grupo Saudável).

Na análise exploratória, a técnica PCA foi utilizada para identificar as características espectrais que apresentaram diferenças entre os grupos Saudável e Câncer, extraindo informações baseadas na variância dos dados. Do conjunto de dados originais foram obtidas duas novas variáveis: os Scores e os PCs (Figura 2). Os Scores correspondem aos eixos de variância espectral, que se assemelham aos espectros Raman e podem



ser correlacionados com constituintes bioquímicos do soro, e os PCs correspondem às intensidades (“concentrações”) que cada Score apresenta nos dados originais. Os Scores podem ser empregados para sugerir a constituição bioquímica das amostras de soro e suas diferenças em função do grupo (Saudável ou Câncer), enquanto os PCs indicam onde (em qual grupo) estão essas diferenças (LASCH, 2012; DA SILVA et al., 2018; LIMA et al., 2022). As oito primeiras variáveis da PCA são apresentadas na Figura 2 e representaram 98,2% da variância total observada no conjunto de dados espectrais (96,2%, 0,6%, 0,4%, 0,4%, 0,2%, 0,2%, 0,1% e 0,1% para Score1 a Score8, respectivamente). As características espectrais apresentadas nos primeiros oito Scores foram associadas aos compostos bioquímicos do soro de acordo com a literatura. O Score com picos positivos e PCs positivos sugere que os constituintes bioquímicos estão com alta concentração no grupo, e o Score com picos positivos e PCs negativos (e vice-versa) sugere que os constituintes bioquímicos estejam em baixa concentração no grupo. O Score1 (Figura 2) demonstrou características espectrais decorrentes dos constituintes do soro sanguíneo, principalmente albumina e globulinas, com picos característicos de albumina em 857, 1453 e 1656 cm^{-1} , semelhantes aos encontrados em espectros Raman de soro humano (GONZÁLEZ-SOLÍS et al., 2014; LIMA et al., 2022; PARACHALIL et al., 2019; LIN; KONIG, 1976; DINGARI et al., 2012), cujas posições dos picos são bem parecidas e estão alta concentração em ambos os grupos estudados. A albumina é o componente mais abundante encontrado no soro de humanos (PARACHALIL et al., 2019) e em animais corresponde a cerca de 35 a 50% do total de proteínas (ROCHA et al., 2010) e está diretamente associado ao equilíbrio de fluidos corpóreos e boa saúde (PARACHALIL et al., 2019; C A alta concentração de albumina pode dificultar a identificação de outros componentes espectrais de concentração menos abundantes no soro sanguíneo (LIMA et al., 2022). O PC1 sugere que as amostras do grupo Câncer possuíam concentração maior de albumina em comparação com as amostras do grupo Câncer, com diferença estatisticamente significativa entre os grupos (Mann-Whitney, $p < 0,001$). A concentração maior de albumina no grupo Câncer encontrada nesse estudo é discordante com os achados dos estudos Raman em soro de pacientes com leucemia (GONZÁLEZ-SOLÍS et al., 2014; DA SILVA et al., 2018) e câncer cervical (GONZÁLEZ-SOLÍS et al., 2014), permitindo utilizar essa



biomolécula para discriminar entre pacientes saudáveis e pacientes com neoplasias. De acordo com Lima et al., (2022), a albumina com níveis séricos inferiores a 3,0 g/dL poderia determinar um fator preditivo de óbito próximo. Estudos apontaram que a baixa concentração de albumina foi associada à baixa sobrevida e taxa de mortalidade em diversos tipos de câncer, independente do escore corporal e da presença da proteína C reativa (PCR) (DO N.S LIMA et al., 2021; DURAN et al., 2014). As posições dos picos atribuídos as globulinas são similares aos da albumina (LIMA et al., 2022; KUHAR et al., 2021), e podem estar sobrepostos aos picos da albumina. O Score 2 apresentou picos negativos em 760, 880, 1004, 1454, 1466 e 1680 cm^{-1} , que podem ser atribuídos a proteínas e/ou aminoácidos: aminoácidos contendo anéis aromáticos em 760 cm^{-1} (triptofano) (GONZÁLEZ-SOLÍS et al., 2014; DA SILVA et al., 2018; DE GELDER et al., 2007; ECKERSALL, BELL, 2010; LIN-VIEN et al., 1991), 880 cm^{-1} (triptofano) (DE GELDER et al., 2007) e 1004 cm^{-1} (fenilalanina)(GOULART et al., 2022; LIMA et al., 2022), e proteínas em 1454 cm^{-1} [δ CH₂ e δ C–N–H] (GOULART et al., 2022; LIMA et al., 2022), 1466 cm^{-1} (δ CH₂/CH₃ (PARACHALIL et al., 2019; KUHAR et al., 2021; DINGARI et al., 2012; SANDRI et al, 2006), e 1680 cm^{-1} (amida I – α -helix) (ROCHA et al., 2010). O PC2 apresentou intensidade próxima de zero e grande desvio padrão, e não apresentou diferença estatisticamente significativa entre os grupos (Mann-Whitney, p = 0.320), indicando que essas proteínas/aminoácidos apresentam grande variação entre os grupos e sugerindo que não estariam associadas com o processo neoplásico, não sendo relevantes para a diferenciação dos grupos estudados. O Score 3 indicou a existência de um “red shift” (deslocamento da banda para frequência de vibração maior) (LIMA et al., 2022; PARACHALIL et al., 2019; KUHAR et al., 2021) dos picos em 621, 642, 1003 e 1032 cm^{-1} (picos invertidos em 617/624, 639/645, 1000/1006, e 1028/1035 cm^{-1}) atribuídos aos aminoácidos que possuem anel aromático: fenilalanina, tirosina e triptofano, que ocorreu no grupo Câncer. Os picos positivos em 1304, 1441 e 1658 cm^{-1} são atribuídos a lipídeos: em 1304 cm^{-1} (δ CH₃/CH₂ (JOLLIFFE, CADIMA, 2016; LIN; KONIG, 1976), em 1441 cm^{-1} (δ CH₃/CH₂) (BANERJEE, ZHANG, 2007) e em 1658 cm^{-1} (δ C=C) (GONZÁLEZ-SOLÍS et al., 2014; LIMA et al., 2022; KRAFT et al., 2005). Os picos negativos em



760, 1132, 1234, 1566 e 1622 cm^{-1} podem ser atribuídos a aminoácidos (triptofano e tirosina) GONZÁLEZ-SOLÍS et al., 2014; LIMA et al., 2022; DA SILVA et al., 2018). O PC3 apresentou diferença estatisticamente significativa entre os grupos (MannWhitney, $p < 0.001$), sugerindo maior concentração de lipídeos no grupo Saudável, e presença de “red shift” e maior concentração de aminoácidos (triptofano e tirosina) no grupo Câncer. Para a fenilalanina, os picos em 621, 1003 e 1032 cm^{-1} estão relacionados a relacionados com os movimentos vibracionais da cadeia lateral e indicam a constância do número de onda ao passar do monómero para o peptídeo e para a proteína (HERNÁNDEZ et al., 2013). As intensidades dos picos em 621 e 1003 cm^{-1} da fenilalanina são sensíveis à hidratação/hidrofobicidade do solvente, devido à interação do solvente com a nuvem de electrões π do anel benzênico substituído (KUHAR et al., 2021), e os picos em 1032 e 1004 cm^{-1} são sensíveis ao pH (devido à protonação) (FISCHER, HANS, 1992), porém sem mudança de posição. O “red shift” pode sugerir uma alteração conformacional de proteínas (dobras das estruturas secundárias e terciárias), ou mesmo ligações ocorrendo em sítios próximos ao anel aromático desses aminoácidos, onde as frequências das bandas das cadeias laterais dos resíduos de aminoácidos são deslocadas para a direita do espectro (“vermelho”) devido a essas ligações (pH e protonação, por exemplo). Apesar da constância do número de onda ao passar do monómero (fenilalanina) ao peptídeo (somatostatina) e à proteína (albumina de soro bovino) vista na Figura 1 do estudo de Hernández et al., (2013) pode-se sugerir maior energia nos picos para as proteínas comparativamente a aminoácidos; apesar de serem muito sutis, essas possíveis mudanças na posição desses picos podem ser detectadas pela alta resolução do espectrógrafo Raman nessa faixa de número da onda (entre 400 e 1000 cm^{-1}). O Score 4 demonstrou picos positivos em 455, 763, 1240, 1361, 1501 e 1529 cm^{-1} que podem ser atribuídos a proteínas e aminoácidos (GONZÁLEZ-SOLÍS et al., 2014; VARGAS-OBIETA et al., 2016; LIN; KONIG, 1976; KUHAR et al., 2021; TALARI et al., 2015). Os picos negativos em 514, 569, 676, 901, 948, 1004, 1304, 1438, 1567 e 1621 cm^{-1} podem ser atribuídos a proteínas, particularmente hemoglobina (BANKAPUR et al., 2010; LIN-VIEN et al., 1991) O PC4 demonstrou diferença estatisticamente significativa entre os grupos (Mann-Whitney, $p < 0.001$), sugerindo maior concentração de proteínas e aminoácidos no grupo Saudável e maior



concentração da hemoglobina no grupo Câncer, o que pode indicar ocorrência de hemólise nas amostras desses grupos. A diminuição de concentração de proteínas/aminoácidos em pacientes oncológicos poderia estar relacionada a um aumento no catabolismo de proteínas musculares, resultando em perda de massa muscular induzida pelo tumor (SANDRI, 2006) e na já mencionada diminuição da albumina nos pacientes oncológicos. A anemia hemolítica por agente oxidante é devido à quebra do metabolismo antioxidativo do eritrócito, formando a metemoglobina, sendo ela a responsável pela desnaturação proteica, ou ainda de origem microangiopatia, onde o eritrócito é fragmentado por alterações da microvasculatura formando trombos (FIGHERA, 2007). No Score 5 é observado o “red shift” do pico 1003 cm^{-1} (picos $1000/1006\text{ cm}^{-1}$), identificado no grupo Câncer (semelhante ao Score 3). Os picos positivos em $756, 1229, 1551, 1562, 1584$ e 1623 cm^{-1} podem ser atribuídos a hemoglobina (FIGHERA, 2007). O PC5 apresentou diferença estatisticamente significativa (t-test, $p < 0.01$), sugerindo maior intensidade de células vermelhas no grupo Saudável, com a já observada ocorrência de anemia no grupo Câncer. O Score 6 também indicou ocorrência de “red shift” do pico 1003 cm^{-1} (picos $1001/1005\text{ cm}^{-1}$) no grupo Câncer. Os picos negativos em $485, 883, 1135, 1227, 1453$ e 1557 cm^{-1} podem ser atribuídos a proteínas (GONZÁLEZ-SOLÍS et al., 2014; LIN; KONIG, 1976; FIGHERA, 2007; NUNES et al., 2012; PARACHALIL et al., 2019; KUHAR et al., 2021; DINGARI et al., 2012; LIN-VIEN et al., 1991). Os picos em 485 e 883 cm^{-1} (LIN-VIEN et al., 1991; BORGES et al., 2015) podem estar atribuídos a PCR. O PC6 não apresentou significância estatística (teste t , $p = 0.8315$), com intensidade próxima de zero e grande desvio padrão, não sendo relevantes para a diferenciação dos grupos. O Score 7 apresentou picos positivos em $883, 1014$ e 1252 cm^{-1} , que podem ser atribuídos a PCR (BERGHOLT, HASSING, 2009), e picos negativos em 945 e 974 cm^{-1} , que podem ser atribuídos a proteínas e aminoácidos. O PC7 apresentou diferença estatisticamente significativa (Mann-Whitney, $p < 0.0001$), sugerindo maior concentração de PCR no grupo Câncer e a já observada maior concentração de proteínas/aminoácidos no grupo Saudável. O Score 8 apresentou picos positivos em $612, 882, 1087, 1268$ e 1441 cm^{-1} , atribuídos a lipídeos (GONZÁLEZ-SOLÍS et al., 2014; LIMA et al., 2022; VARGAS-OBIETA et al., 2016; KRAFT et al., 2005; TALARI et al., 2015) e picos negativos em $511,$



900, 950, 1005, 1211, 1343 e 1416 cm^{-1} , atribuídos a proteínas/aminoácidos (GONZÁLEZ-SOLÍS et 2014; DA SILVA et al., 2018; LIN-VIEN et al., 1991; ATKINS et al., 2017; PICHARDO-MOLINA et al., 2007). O PC8 não apresentou diferença estatisticamente significativa (*t*-test with Welch correction, $p = 0.3522$), com intensidades próximas de zero e grande desvio padrão, indicando que esses picos não são relevantes para a diferenciação dos grupos. Os resultados revelaram que as maiores diferenças foram observadas nos picos em 1132, 1342, 1368 e 1453 cm^{-1} , que sugerem concentrações maiores de proteínas e/ou aminoácidos no grupo Câncer, e picos em 451 cm^{-1} (triptofano) (GONZÁLEZ-SOLÍS et 2014; DE GELDER et al., 2007) e 1441 cm^{-1} (lipídeos) (GONZÁLEZ-SOLÍS et 2014; VARGAS-OBIETA et al., 2016; KRAFT et al., 2005), que sugerem concentrações maiores no grupo Saudável, e a ocorrência de “red shift” em 617/624, 1000/1006 e 1028/1034 cm^{-1} no grupo Câncer. Os picos negativos/positivos em 617/624, 1000/1006 e 1028/1034 cm^{-1} sugerem “red shift” dos picos em 621, 1003 e 1032 cm^{-1} (fenilalanina) (GONZÁLEZ-SOLÍS et 2014; MARTINEZ-ESPINOSA et al., 2008; DE GELDER et al., 2007; HERNÁNDEZ et al., 2013) está presente no espectro do grupo Câncer. Lima et al. (2022) verificaram nos espectros Raman de soro de pacientes infantojuvenis que proteínas (principalmente albumina), aminoácidos e carotenoides estavam em intensidade maior no grupo normal comparativamente aos grupos leucêmico e outros tipos de câncer, discordando dos resultados obtidos neste estudo principalmente com relação à albumina. A obtenção do proteoma sérico com a identificação de proteínas de fase aguda em inflamação é uma ferramenta importante em exames de diagnóstico veterinário (ATHERTON et al., 2013), pois as alterações na composição e concentrações de proteínas plasmáticas podem ocorrer em linfoma canino comparativamente ao indivíduo saudável, e valores aumentados de proteínas são comumente observados em inflamação decorrente do câncer (ECKERSALL, BELL, 2010). Picos de carotenoides não foram observados na análise PCA, sugerindo não estarem presentes na dieta desses animais. Em estudos por espectroscopia Raman com soro e plasma de seres humanos, Silva et al. (2018) e Gonzáles-Solís et al. (2014) analisaram espectros Raman de plasma humano de pacientes leucêmicos e saudáveis. González-Solís et al. (2014) analisaram espectros de soro sanguíneo humano de pacientes com câncer cervical, pré-câncer e saudáveis,



e esses estudos identificaram picos de carotenoides significativamente mais intensos nos indivíduos normais comparativamente a indivíduos neoplásicos. Munteanu et al. (2017) analisaram os espectros Raman (SERS) de dois fragmentos de tecidos mamários benignos e malignos em cães e demonstraram a presença de carotenoides exclusivamente em tecidos benignos.

Análise Discriminante por PLS. Técnicas de discriminação baseadas em PLS e PCA (PLS-DA e PCA-DA) tem sido usada com sucesso como preditores para discriminação ou diferenciação de amostras nas diversas áreas do conhecimento (GOHEEN et al., 1977; SILVEIRA et al, 2012), particularmente a PLS-DA. A PLS-DA possibilita conhecer as variâncias dentro dos grupos estudados e classificar grupos amostrais baseada nas diferenças existentes entre grupos aplicada na área oncológica (LIMA et al., 2022; SILVA et al., 2018; NGUYEN, ROCKE, 2002; PICHARDO-MOLINA et al, 2007) O modelo PLS-DA foi aplicado aos espectros de soro dos grupos Saudável e Câncer com validação cruzada “leave one out”, onde o número de variáveis latentes que melhor descreve o modelo é aquela que promove a maior acurácia de classificação. Os resultados da PLS-DA usando a validação cruzada com 7 LVs são apresentados na Tabela 4, e demonstraram que a classificação dos espectros de soro dos pacientes nos grupos analisados (Saudável x Câncer) interpretados via “spectrum wise” apresentou acurácia de 77.1%, sensibilidade de 77.0% e especificidade de 77.3%, enquanto que a discriminação obtida via “patient wise” apresentou acurácia de 77.7%, sensibilidade de 75.6%, e especificidade 80.7%; a interpretação “patient wise” demonstrou melhor diferenciação. Gonzáles-Solís et al. (2014) utilizaram a espectroscopia Raman no soro humano para monitorar pacientes leucêmicos que estavam recebendo quimioterapia, através da PCA foram identificados diferenças referentes aos componentes proteicos, aminoácidos e carotenóides que estavam presentes no grupo leucêmico, obtendo 100% de sensibilidade e especificidade através da validação cruzada “leave-one-out” (KUHAR et al., 2021; SILVA et al. 2018) utilizaram a espectroscopia Raman e análise discriminante por PLS e PCA buscando diagnóstico de leucemia em plasma humano, que resultou em uma acurácia de 97,1%, sensibilidade de 95,7% e especificidade de 98,0% para a PLS. Lima et al. (2022) utilizaram a espectroscopia Raman no soro de crianças e adolescentes para

discriminar amostras normais das amostras de câncer leucêmicas e não leucêmicas, utilizando PCA e PLS-DA, resultando para modelo de classificação dos espectros a acurácia de 66%, sensibilidade 99% e especificidade 57%. Uma vantagem da utilização da PLS-DA em relação à PCA-DA é a que a primeira busca maximização da covariância entre as variáveis espectrais e os grupos que as amostras pertencem, apontando um melhor desempenho na classificação de amostras discriminando-as quando comparadas com a PCA (NGUYEN, ROCKE, 2002), que busca a variância total dos dados. A espectroscopia Raman se destaca por ser realizada de forma minimamente invasiva, sem preparo específico das amostras como corantes ou reagentes, e utilizando apenas uma pequena quantidade de material biológico (HANLON et al., 2000), com potencial para ser utilizada na Medicina Veterinária (DUARTE et al., 2010; MUNTEANU et al., 2017; BIRTOIU et al., 2016). Uma vantagem na espectroscopia Raman em doenças neoplásicas é sugerir um diagnóstico com relativa precocidade (SILVEIRA et al., 2012; MELITTO et al. BIRTOIU et al., 2016; DUKOR, 2006). Contudo, não poderia ainda competir com a biópsia anatomopatológica (padrão ouro) (BRAZ et al., 2016) que fornece uma acurácia superior a 90%. É necessário realizar mais estudos com animais em situações com maior controle de alimentos e doenças pré-existentes para modelar e reduzir a quantidade de erros relacionados à classificação dos pacientes, a fim de ampliar os dados relacionados à fisiologia de animais saudáveis e doentes. Recomenda-se que os testes de diagnóstico sejam realizados concomitantemente à anamnese e semiologia maximizando as chances de sucesso no diagnóstico e tratamento, pois apenas a citologia aspirativa (BRAZ et al., 2016), que possui 69,69% de acurácia em comparação à histopatologia, poderia ser um método que pudesse concorrer com a espectroscopia Raman. O presente estudo demonstrou que a espectroscopia Raman aplicada no soro de cães tem potencial para ser utilizada como auxiliar no diagnóstico precedendo a confirmação por exame anatomopatológico, e que esta técnica poderia oferecer precocidade e com isso otimizar o tempo de início dos protocolos terapêuticos. Este método óptico poderia ser usado como uma análise complementar no rastreio precoce de alterações bioquímicas de modo auxiliar o



diagnóstico diferencial nos grupos estudados (MARTINEZ-ESPINOSA et al., 2008; LIMA et al., 2022).

Conclusões:

Neste estudo preliminar, a espectroscopia Raman combinada às técnicas de análise estatística multivariada PCA e PLS foram utilizadas para identificar alterações bioquímicas em amostras de soro de cães com câncer comparativamente ao de cães clinicamente saudáveis principalmente as proteínas, aminoácidos, lipídeos e utilizar essas diferenças para classificar as amostras em Saudável e Câncer. O espectro da diferença revelou picos de proteínas e aminoácidos, especialmente à albumina e fenilalanina, que estavam em maior intensidade no grupo Câncer, e pico de lipídeos e de triptofano, que estavam em maior intensidade no grupo Saudável. Na análise exploratória via PCA, foi observado no grupo Câncer aumento de proteínas/aminoácidos (Score1/PC1 e Score3/PC3), redução de lipídeos (Score3/PC3 e Score7/PC7), aumento de red blood cells (Score 5/PC5) e aumento da PCR e da glucosamina (Score 7/PC7). Foi observado o “red shift” no grupo Câncer. A PLS-DA aplicada aos espectros Raman permitiu a classificação dos espectros nos grupos (Saudável e Câncer) utilizando a validação cruzada “leave-one-out” e classificação dos espectros na forma “patient wise” atingiu acurácia de 77.7%, com sensibilidade de 75.6% e especificidade de 80%. A espectroscopia Raman constitui-se como uma ferramenta específica e sensível, minimamente invasiva e sem a necessidade de preparo da amostra, que pode vir a ser utilizada no diagnóstico complementar à biópsia histopatológica, tendo potencial para a classificação dos animais saudáveis e portadores de câncer através das diferenças espectrais que se relacionam com as alterações bioquímicas presentes no soro.

REFERÊNCIAS:

SCHIFFMAN J.D., M. BREEN, Comparative oncology: what dogs and other species can teach us about humans with cancer. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, v. 370, n.1673. (2015).

GONZÁLEZ-SOLÍS J.L., MARTÍNEZ-ESPINOSA, J.C; TORRES-GONZÁLEZ L.A.; AGUILAR-LEMARROY, A.; JAVE-SUÁREZ, L.F.; PALOMARES-ANDA P. Cervical cancer detection based on serum sample Raman spectroscopy, *Lasers Medical Science.*, v.29, n.3, 979-85. (2014)



SILVEIRA, L.; SILVEIRA, F.L.; BODANESE, B.; ZÂNGARO, R.A.;
PACHECO, M.T.T. *Journal of Biomedical Optics*, v. 17, n.7. (2012)

ECKERSALL, P. D; BELL, R. Acute phase proteins: Biomarkers of infection and inflammation in veterinary medicine. *The Veterinary Journal*, v.185, n.1, 23-7. (2010)

DUKOR, R.K. Handbook of Vibrational Spectroscopy, Wiley, Downers Grove, IL, USA. (2006).

FOMENTO

D. F. Lopes agradece ao Instituto Doutor Izaio Soares - IDIS pelo apoio financeiro. L. Silveira Jr. agradece à Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) pela concessão do espectrômetro Raman (Processo No. 2009/01788-5) e ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pela bolsa de produtividade (Processo No. 314167/2021-8).

