

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
Faculdade de Farmácia
Disciplina de Estágio Curricular em Farmácia

**Análise estrutural de saponinas com atividade imunoadjuvante e hemolítica
mediante uso de ferramentas estatísticas**

Samuel Kaiser

Porto Alegre, Junho de 2008.

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
Faculdade de Farmácia
Disciplina de Estágio Curricular em Farmácia

**Análise estrutural de saponinas com atividade imunoadjuvante e hemolítica
mediante uso de ferramentas estatísticas**

Samuel Kaiser
Trabalho de Conclusão
da Disciplina de Estágio Curricular em Farmácia

Prof. Dr. George González Ortega
Orientador

Prof. Dr. Hugo Verli e M.Sc Cabral Pavei
Co-orientadores

Porto Alegre, Junho de 2008.

**Às pessoas mais importantes
da minha vida.
Meus pais:
Iara Maria Hatwig Kaiser
e Leoni Kaiser
Meu avô e maior incentivador:
Afonso Guilherme Hatwig (*in memoriam*)**

**Não vim até aqui para desistir agora,
se depender de mim
eu vou até o fim.**

Humberto Gessinger

AGRADECIMENTOS

Aos colegas do Laboratório de Desenvolvimento Galênico (LDG) da Faculdade de Farmácia da UFRGS, em especial ao doutorando Cabral Pavei pela colaboração prestada e todo o apoio para a realização do trabalho.

Ao Professor Dr. Hugo Verli por ter me recebido muito bem no Laboratório de Bioinformática Estrutural e todo apoio e incentivo para a realização do trabalho.

Ao Professor Dr. George González Ortega pelo exemplo de profissional e pelo seu enorme interesse e preocupação com o aprendizado do aluno e também pela orientação prestada na realização do trabalho.

Aos meus familiares, em especial aos meus pais Leoni Kaiser e Iara Maria Hatwig Kaiser, pelo apoio e amor incondicional em todas as etapas de minha vida.

Ao meu avô e grande incentivador Afonso Guilherme Hatwig (*in memoriam*), por todos os conselhos que me levaram à Faculdade de Farmácia e que me motivam, até hoje, a sempre “seguir em frente” e, portanto, influenciaram e muito na realização deste trabalho a ele dedicado.

SUMÁRIO

	LISTA DE TABELAS.....	I
	LISTA DE FIGURAS.....	II
	GLOSSÁRIO.....	IV
	RESUMO.....	VII
1	INTRODUÇÃO.....	1
2	OBJETIVOS.....	7
3	METODOLOGIA.....	9
3.1	Análise estrutural.....	10
3.1.1	Seleção das saponinas.....	10
3.1.2	Avaliação dos grupamentos substituintes e atividades biológicas.....	10
3.1.3	Análise de agrupamentos.....	11
3.1.4	Análise de fatores.....	11
3.1.4.1	Análise de componentes principais.....	11
3.1.4.2	Análise de fatores comuns.....	12
3.1.5	Análise da capacidade preditiva do modelo proposto.....	13
4	RESULTADOS E DISCUSSÃO.....	14
4.1	Análise estrutural.....	15
4.1.1	Seleção das saponinas.....	15
4.1.2	Avaliação dos grupamentos substituintes e atividades biológicas.....	17
4.1.3	Análise de agrupamentos.....	21
4.1.4	Análise de fatores.....	28
4.1.4.1	Análise de componentes principais.....	28
4.1.4.2	Análise de fatores comuns.....	38
4.1.5	Análise da capacidade preditiva do modelo proposto.....	41
5	CONCLUSÕES.....	44
6	PERSPECTIVAS.....	47
7	REFERÊNCIAS.....	49

LISTA DE TABELAS

Tabela 1.	Representação estrutural das saponinas triterpênicas derivadas de (A) <i>G. max</i> . (B) <i>D. lablab</i> e (C) <i>A. hippocastanum</i>16
Tabela 2.	Grupamentos substituintes e abreviações das variáveis utilizadas na análise estrutural das saponinas triterpênicas.....17
Tabela 3.	Codificação das variáveis referentes aos substituintes e às atividades imunoadjuvante e hemolítica das saponinas triterpênicas provenientes de <i>A. hippocastanum</i> , <i>D. lablab</i> e <i>G.max</i>19 e 20
Tabela 4.	Representação estrutural das saponinas triterpênicas utilizadas na avaliação da capacidade preditiva do modelo proposto.....42

LISTA DE FIGURAS

Figura 1.	Estrutura-modelo de uma saponina triterpênica	2
Figura 2.	Diagrama da metodologia aplicada à análise estrutural.....	12
Figura 3.	Dendrogramas das observações e variáveis na (A e B) ausência e na (C e D) presença de atividades biológicas codificadas.....	22
Figura 4.	Dendrogramas das observações e variáveis na (A e B) presença de atividades biológicas codificadas e na (C e D) presença de atividades biológicas representadas por variáveis numéricas.....	25
Figura 5.	Gráficos de autovalor das observações e variáveis na (A e B) ausência e na (C e D) presença de atividades biológicas codificadas.....	30
Figura 6.	Representação gráfica da análise de componentes principais das observações e variáveis na (A e B) ausência e (C e D) na presença de atividades biológicas codificadas.....	31
Figura 7.	Gráficos de autovalor das observações e variáveis na (A e B) presença de atividades biológicas codificadas e na (C e D) presença de atividades biológicas representadas por variáveis numéricas.....	34
Figura 8.	Representação gráfica da análise de componentes principais	

das observações e variáveis na (**A e B**) presença de atividades biológicas codificadas e na (**C e D**) presença de atividades biológicas representadas por variáveis numéricas.....35

Figura 9. Gráficos de autovalor das observações na (**A**) presença de atividades biológicas codificadas e na (**B**) presença de atividades biológicas representadas por variáveis numéricas.....39

Figura 10. Representação gráfica da análise de fatores comuns das observações na (**A e C**) presença de atividades biológicas codificadas e na (**B e D**) presença de atividades biológicas representadas por variáveis numéricas , sendo que, **A e B** representam a matriz não-rotacionada e **C e D** representam a matriz rotacionada.....40

GLOSSÁRIO

Agrupamento: o conjunto de variáveis que representam as características usadas para comparar observações na técnica de análise de agrupamentos.

Análise de agrupamentos: técnica multivariada que permite agrupar observações ou variáveis com perfis similares em um dado conjunto de características.

Análise de componentes principais: modelo fatorial no qual os fatores são baseados na variância total. Na análise de componentes toda a variância é comum ou compartilhada.

Análise de fatores: nome genérico dado a uma classe de métodos estatísticos multivariados (análise de componentes principais e fatores comuns) cujo propósito principal é definir a estrutura subjacente em uma matriz de dados.

Análise de fatores comuns: modelo fatorial no qual os fatores são baseados em uma matriz de correlação reduzida. Na análise fatorial os fatores extraídos são baseados apenas na *variância comum*, com as *variâncias específicas* e de *erro* excluídas.

Análise multivariada: são todos os métodos estatísticos que simultaneamente analisam múltiplas medidas de observações sob investigação. Qualquer análise simultânea de mais de duas variáveis pode, de certo modo, ser considerada uma análise multivariada.

Cargas fatoriais: correlação entre as variáveis originais e os fatores, bem como a chave para o entendimento da natureza de um fator em particular.

Componente: mesmo significado de *fator*, diferindo deste por ser aplicável à análise de componentes principais, enquanto o termo *fator* é utilizado na análise de fatores comuns.

Critério da raiz latente: técnica mais utilizada na análise de componentes principais e fatores comuns para a determinação do número de fatores a serem extraídos.

Fator: combinação linear (variável estatística) das variáveis originais. Os fatores também representam as dimensões latentes (constructos) que resumem ou explicam o conjunto original de variáveis observadas.

Ligação individual: é um dos métodos que compõem o procedimento hierárquico aglomerativo baseado nas distâncias mínimas, ou seja, o primeiro agrupamento é formado por duas observações separadas pela menor distância (maior similaridade), e assim sucessivamente até que todas as observações formem um só agrupamento.

Matriz de correlação: tabela que mostra as intercorrelações entre todas as variáveis.

Matriz de origem: matriz na qual as observações representam as linhas e as variáveis as colunas.

Matriz transposta: matriz na qual as observações representam as colunas e as variáveis as linhas.

Matriz rotacionada: matriz submetida ao método de rotação ortogonal.

Matriz não-rotacionada: matriz não submetida ao método de rotação ortogonal.

Medidas do tipo correlacionadas: representam a similaridade pela correspondência de padrões (observações) ao longo das características (variáveis).

Método rotacional ortogonal: método pelo qual os fatores são rotacionados em torno da origem, mantendo sempre um ângulo de 90° um em relação ao outro, até que alguma outra posição seja alcançada.

Nível de similaridade: correspondência ou associação baseada nas variáveis da variável estatística de agrupamento.

Procedimento hierárquico aglomerativo: procedimento de agrupamento que envolve uma combinação das observações em agrupamentos. O resultado é a construção de uma hierarquia, ou estrutura em árvore (dendrograma) que representa a formação dos agrupamentos. Tal procedimento produz $N - 1$ soluções de agrupamentos, onde N é o número de observações

Regra de parada: algoritmo para determinar o número final de agrupamentos a serem formados. Existem duas classes de regras que são aplicadas *post hoc* e calculadas sendo elas compostas pelas medidas de similaridade e medidas estatísticas adaptadas.

Variância comum: variância compartilhada com outras variáveis na análise fatorial.

Variância do erro: variância de uma variável devido a erros de coleta de dados ou na medida.

Variância específica: variância de cada variável, única àquela variável que não é explicada ou associada com outras variáveis na análise de fatores.

Variável dicotômica: variável métrica binária usada para representar uma única categoria de uma variável não-métrica.

VARIMAX: é um dos métodos de rotação fatorial ortogonal mais amplamente aplicado.

RESUMO

As saponinas são metabólitos secundários presentes em diversas espécies, como, por exemplo, *Glycine max.*, *Quillaja saponaria* e *Panax ginseng*, entre outras, que apresentam importantes atividades físico-químicas e farmacológicas, dentre as quais se destacam suas atividades imunoadjuvante e hemolítica. O aumento do número de saponinas avaliadas quanto ao seu potencial imunoadjuvante e hemolítico associada à ausência de trabalhos de sistematização das informações disponíveis, levaram ao estabelecimento de relações estrutura-atividade vagas e pouco conclusivas para essa classe de substâncias. Neste sentido, o presente trabalho visa avaliar as relações estruturais envolvidas com as atividades imunoadjuvante e hemolítica das saponinas triterpênicas selecionadas, mediante análise estrutural via técnicas estatísticas multivariadas. Para isso, foram selecionadas 20 saponinas de estrutura triterpênica derivadas, principalmente, do ácido oleanólico, cujos potenciais imunoadjuvante e hemolítico foram relatados na literatura. Essas saponinas foram analisadas, posteriormente, em relação aos grupamentos substituintes do anel triterpênico e codificadas na forma de *variáveis dicotômicas*. As atividades biológicas foram adicionadas à análise, tanto na forma de *variáveis dicotômicas* de caráter qualitativo, como na forma de variáveis numéricas de caráter quantitativo. Estas codificações deram origem uma matriz de dados que foi submetida à *análise de agrupamentos*, *análise de componentes principais* e *análise de fatores comuns*. Os resultados obtidos revelam que os métodos aplicados são sensíveis, seletivos e específicos, sendo que a atividade hemolítica das saponinas parece estar relacionada à presença de uma hidroxila em C-16, de CH₂OH em C-17, de uma acetila em C-22 e de um grupamento acila em C-21. Já a atividade imunoadjuvante pode estar relacionada à presença de uma hidroxila em C-21, de uma metila em C-17, e de dois açúcares ou cadeia de açúcares ligados à aglicona. Além disso, foi observada uma relação inversa entre as atividades hemolítica e imunoadjuvante, indicando que uma saponina que apresenta um alto potencial imunoadjuvante tende a apresentar um baixo potencial hemolítico e vice-versa.

1 INTRODUÇÃO

1 INTRODUÇÃO

As saponinas são heterosídeos, provenientes principalmente de espécies vegetais, constituídos por um núcleo esteróide ou triterpênico, denominado aglicona, ao qual estão ligadas uma ou mais cadeias de açúcares ou osídicas. As saponinas triterpênicas apresentam um núcleo fundamental composto por cinco anéis e 30 átomos de carbono, ao passo que as saponinas esteróides possuem a aglicona constituída por 27 carbonos. Com relação à porção osídica, as saponinas são denominadas monodesmosídicas quando a aglicona apresenta-se substituída por apenas um açúcar ou uma cadeia de açucares, enquanto que as saponinas bidesmosídicas apresentam dois açúcares ou duas cadeias de açucares ligadas à aglicona (figura 1) (HOSTETTMANN e MARSTON, 1995; SHENCKEL e col., 2003).

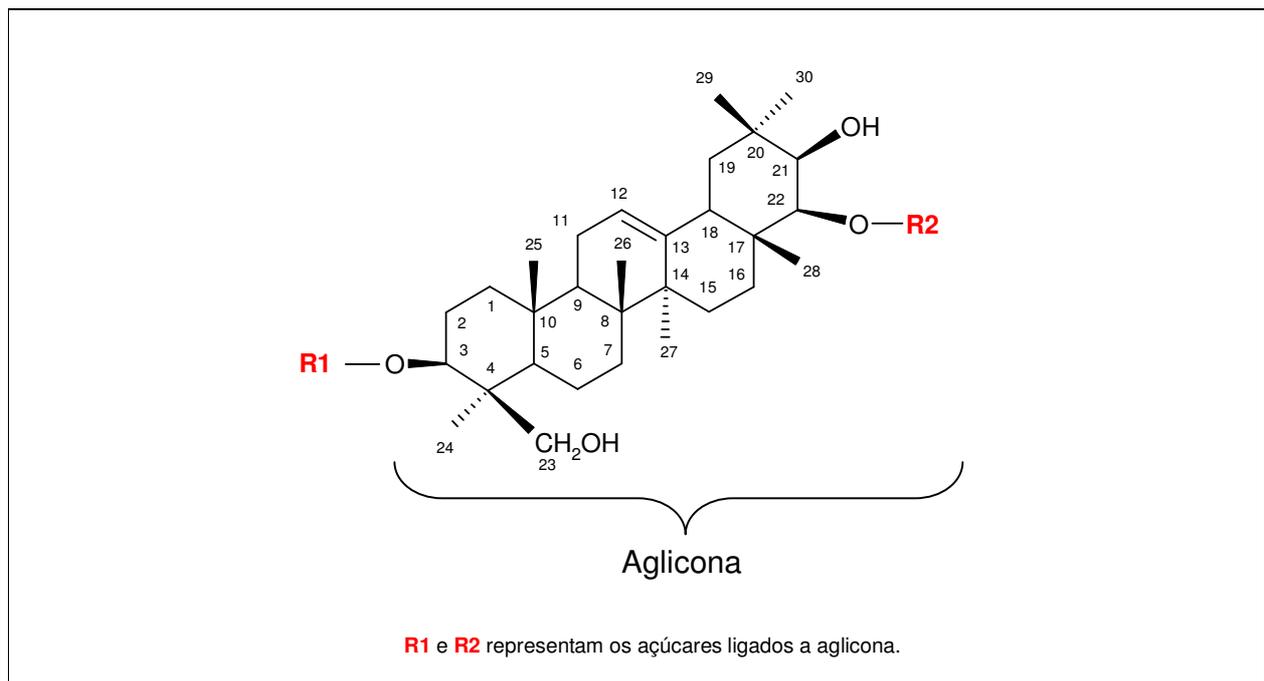


Figura 1. Estrutura-modelo de uma saponina triterpênica.

Dentre os metabólitos secundários produzidos pelos vegetais, as saponinas constituem uma das classes de maior destaque devido à sua ampla ocorrência no reino vegetal e às suas importantes atividades farmacológicas e físico-químicas. No que

tange aos aspectos farmacológicos, as saponinas apresentam importantes atividades, como por exemplo, antiinflamatória, hipocolesterolêmica, hemolítica, antiviral, imunoestimulante e imunoadjuvante. (WANG e NG, 1999; ODA e col., 2000; SÜLEYMAN e col., 2003; SCHENKEL e col.; 2003), além das atividades relacionadas com as suas propriedades de interface como, por exemplo, a diminuição da tensão superficial. Algumas espécies têm merecido especial destaque no âmbito farmacêutico devido aos elevados potenciais imunoadjuvantes, como *Glycine max*, *Quillaja saponaria*, *Panax ginseng*, *Dolichos lablab*, entre outras.

A atividade imunoadjuvante das saponinas tem chamado a atenção em virtude do surgimento de novas gerações de vacinas, particularmente aquelas baseadas em proteínas recombinantes e DNA, que são menos reativas e imunogênicas se comparadas às vacinas tradicionais. Esta necessidade tem acelerado a busca de novos e efetivos adjuvantes (O'HAGAN e col., 2001; SUN, 2006). Embora uma série de compostos tenham sido testados em vacinas experimentais, estes em sua maioria, apresentaram baixa capacidade imunogênica ou importantes efeitos adversos que limitam a sua aplicação como adjuvantes em vacinas (AUCOUTURIER e col., 2001; HUNTER, 2002; SUN, 2006).

A propriedade imunoadjuvante das saponinas foi descoberta em 1920 e a sua utilização como adjuvante em vacinas para animais foi descrita pela primeira vez em 1951 (OLESZEK e MARSTON, 2000). Posteriormente, diversos estudos foram realizados utilizando saponinas como adjuvantes em vacinas de uso veterinário, principalmente as provenientes de *Q. saponaria*. Isso levou a uma crescente investigação da atividade imunoadjuvante em outras espécies vegetais, entre as quais destacam-se a soja (*G. max*) e ginseng (*P. ginseng*). Os resultados dessas avaliações demonstraram uma atividade promissora das saponinas dessas espécies como adjuvantes em vacinas (ODA e col., 2003; SUN e col., 2007).

A maior dificuldade para o desenvolvimento de uma nova tecnologia em vacinas é encontrar adjuvantes principalmente sintéticos que tenham uma boa segurança e imunogenicidade, ou seja, que sejam capazes de aumentar a resposta imune a um determinado antígeno sem efeitos adversos significativos. Poucos foram os adjuvantes avaliados até o momento em vacinas para humanos, sendo que os sais de alumínio

foram os mais utilizados, embora apresentem baixa atividade imunoadjuvante (OLESZEK e MARSTON, 2000).

Parte da teoria, segundo a qual grupamentos particulares ligados à aglicona, como aldeído, carboxila e acila, seriam determinantes para a atividade imunoadjuvante, parece não encontrar sustentação na atualidade, uma vez que saponinas desprovidas dessas características também apresentaram capacidade imunoadjuvante. Alguns pesquisadores reconhecem a possibilidade dessa atividade estar relacionada com o tamanho e a natureza das cadeias osídicas laterais, contudo, as evidências experimentais a esse respeito ainda são insuficientes. Isto se deve à dificuldade de se obter compostos análogos com atividade imunoadjuvante para investigação. Além disso, muitas informações relacionadas à atividade imunoadjuvante se referem especificamente às saponinas de *Quillaja* ou seus derivados semi-sintéticos, os quais, embora apresentem uma excelente atividade imunoadjuvante, são estruturalmente diferentes de outras saponinas com atividade imunoadjuvante relatada. O grau de pureza das saponinas testadas em diferentes trabalhos também compromete a exatidão dos resultados obtidos em estudos de resposta imunológica, impondo dificuldades adicionais ao estabelecimento de relações estrutura-atividade concisas e conclusivas (ODA e col., 2003).

Entre as atividades farmacológicas, a propriedade hemolítica das saponinas é considerada um efeito negativo dessa classe de substâncias, a qual tem limitado a sua utilização como adjuvante em vacinas. Vários trabalhos sobre a atividade imunoadjuvante e hemolítica foram realizados, na procura de saponinas com alta capacidade imunogênica e baixa atividade hemolítica, contudo, em nenhum deles é exposta uma hipótese ou teoria geral que estabeleça parâmetros de diferenciação ou de correlação entre ambas as atividades (ESTRADA e col., 2000; SUN e col., 2005; SUN e PAN, 2006; SUN, 2006; FLECK e col., 2006). De modo geral, se pode constatar que os trabalhos relacionados à atividade imunoadjuvante e hemolítica de saponinas disponíveis na literatura, buscam em sua maioria estabelecer relações estrutura-atividade. Contudo, este tipo de abordagem tem como base apenas a análise estrutural focada, principalmente, na ausência ou presença de determinados grupamentos relacionando-os com uma maior ou menor atividade. O resultado disso tem sido o

estabelecimento de relações estrutura-atividade, muitas vezes, pouco claras e conclusivas.

Tendo como base os resultados dos estudos realizados até o momento, não há uma relação direta estabelecida entre as atividades hemolítica e imunoadjuvante (ODA e col., 2000). Existem fortes indícios de que a atividade hemolítica das saponinas esteja relacionada ao tipo de aglicona presente em sua estrutura, sendo que as esteróides apresentam uma maior atividade hemolítica se comparadas a triterpênicas (TAKECHI e TANAKA, 1995). Essa maior atividade hemolítica das saponinas esteróides pode estar relacionada à sua capacidade de interagir com maior intensidade com o colesterol das membranas celulares (GLAUERT e col., 1962; KARABALIEV e col., 2003). Também foi demonstrado que a propriedade hemolítica das saponinas não está relacionada com sua tensão superficial (STEURER, 1999). Já a atividade imunoadjuvante das saponinas depende da sua característica anfifílica e, no caso das saponinas de *Quillaja*, dos grupamentos presentes em sua estrutura, como aldeído, carboxila, acila ou a porção osídica ligada a aglicona (ASCARATEIL e DUPUIS, 2006; SUN e col., 2007). Além disso, a atividade imunoadjuvante pode estar relacionada com a conformação molecular das saponinas que depende, indiretamente, dos grupamentos substituintes e dos açúcares ligadas a aglicona (SUN e col., 2007).

Um dos motivos do insucesso no estabelecimento de uma correlação estrutura-atividade das saponinas é a ausência de trabalhos de sistematização das informações disponíveis. Nesse sentido algumas ferramentas estatísticas parecem representar uma alternativa viável.

A *análise de fatores* é um método estatístico que faz parte da *análise multivariada* de dados, que tem como objetivo, resumir informações contidas em diversas variáveis originais em um conjunto menor de novas dimensões, denominado *fator*. Esta técnica permite estabelecer correlações entre as diversas variáveis, que fazem parte de uma matriz complexa, propiciando o aparecimento de relações entre as variáveis que de outro modo não se tornam claras. Além disso, este método permite estimar mediante as denominadas *cargas fatoriais* a importância ou a contribuição que uma determinada variável tem sobre um determinado *fator*, propiciando uma idéia de

quais as variáveis são as mais significativas dentro de todo o conjunto que compõe o *fator* (HAIR e col., 2005).

A *análise de agrupamentos* ou *cluster*, assim como a *análise de fatores*, faz parte da *análise multivariada* de dados. Esta ferramenta estatística permite agregar objetos heterogêneos entre si segundo determinadas características, de modo a conseguir agrupamentos distintos, mas com elevada homogeneidade interna. Ao contrário da *análise de fatores*, que visa correlacionar as variáveis que caracterizam os objetos, a *análise de agrupamentos* visa, principalmente, agregar os objetos com base em suas características que são determinadas pelas variáveis. Portanto, a *análise de fatores* e a *análise de agrupamentos*, possuem um papel complementar, podendo ser muito úteis quando se deseja determinar dimensões latentes, ou seja, que não são factíveis de serem facilmente observadas. (HAIR e col., 2005).

2 OBJETIVOS

2 OBJETIVOS

Objetivo geral:

O presente trabalho visa avaliar as relações estruturais envolvidas com as atividades imunoadjuvante e hemolítica das saponinas triterpênicas selecionadas, mediante análise estrutural via técnicas estatísticas multivariadas.

Objetivos específicos:

- Determinar as relações estruturais envolvidas com as atividades imunoadjuvante e hemolítica das saponinas triterpênicas selecionadas;
- Avaliar a aplicabilidade da *análise de agrupamentos e de fatores* na análise estrutural das saponinas triterpênicas selecionadas;
- Desenvolver um modelo que abranja um conjunto finito de saponinas triterpênicas, no qual seja possível prever a potencialidade de determinada saponina em relação às atividades imunoadjuvante e hemolítica.

3 METODOLOGIA

3 METODOLOGIA

3.1 Análise estrutural

3.1.1 Seleção das saponinas

Foram selecionadas 20 saponinas de estrutura triterpênica, isoladas das espécies *Aesculus hippocastanum* (L.) Merrill, *Dolichos lablab* (L.) Merrill e *Glycine max* (L.) Merrill.

3.1.2 Avaliação dos grupamentos substituintes e atividades biológicas

As saponinas foram avaliadas estruturalmente considerando os substituintes e os açúcares ligados à aglicona nas posições C-3, C-4, C-16, C-17, C-21, C-22, C-28, sendo codificadas na forma de *variáveis dicotômicas* (1 indica presença e 0 ausência de determinado grupamento substituinte ou açúcar). Os resultados referentes às atividades imunoadjuvante e hemolítica das saponinas foram extraídos do estudo realizado por ODA e col. (2000) e expressas pelo Título de Hemaglutinação Passiva (**PHA**) e da Concentração Hemolítica 50% (**CH50%**), respectivamente. Estas variáveis numéricas foram transformadas mediante aplicação de uma escala numérica de 0,5 a 4,5, para **PHA**, e de 0,5 a 2 para **CH50%**. Estas atividades biológicas também foram codificadas na forma de *variáveis dicotômicas*, onde a atividade imunoadjuvante (representada pela variável **Adj**) foi considerada presente quando o **PHA** da saponina era maior que o do branco. Já a atividade hemolítica (representada pela variável **Hemo**) foi considerada presente quando a **CH50%** era menor que 600 µg/ml.

OBS: O teste de Hemaglutinação Passiva, que dá origem ao **PHA**, é um dos métodos mais simples para a quantificação de anticorpos e antígenos. Utilizam-se hemácias de carneiro ou hemácias humanas como elemento revelador. Os antígenos se fixam à superfície das hemácias, tornando-as susceptíveis a hemaglutinação na

presença de um anticorpo. Os protocolos para avaliação tanto da atividade imunoadjuvante via **PHA**, como da atividade hemolítica via **CH50%**, estão descritos detalhadamente no estudo de ODA e col. (2000).

3.1.3 Análise de agrupamentos

A *análise de agrupamentos* foi realizada com auxílio do programa estatístico Minitab 14®. O *procedimento hierárquico aglomerativo* utilizado foi o de *ligação individual* e as medidas utilizadas foram do tipo *correlacionadas* (correlação dos coeficientes de distância). O número de agrupamentos a serem extraídos foi direcionado pela diferença entre o *nível de similaridade* do novo *agrupamento* formado a cada passo, em relação ao anterior. Foram realizadas análises da matriz de origem e da matriz transposta, na ausência de atividades biológicas, na presença de atividades biológicas codificadas e na presença de atividades biológicas representadas por variáveis numéricas, **PHA** e **CH50%**.

3.1.4 Análise de fatores

A análise de fatores foi realizada com auxílio do programa estatístico Minitab 14® sendo utilizada tanto a *análise de componentes principais* como *análise de fatores comuns*.

3.1.4.1 Análise de componentes principais

A *matriz de correlação* foi utilizada na análise, sendo o número de *componentes* extraídos determinado mediante o *critério da raiz latente*. Foram realizadas análises da *matriz de origem* e da *matriz transposta*, na ausência de atividades biológicas, na presença de atividades biológicas codificadas e na presença de atividades biológicas representadas por variáveis numéricas, **PHA** e **CH50%**.

3.1.4.1. Análise de fatores comuns

O número de *fatores* extraídos foi determinado pelo *critério da raiz latente*. A rotação dos fatores foi realizada pelo *método rotacional ortogonal, VARIMAX*. Foi realizada a análise da matriz transposta *rotacionada e não-rotacionada*, na presença de atividades biológicas codificadas e na presença de atividades biológicas representadas por variáveis numéricas, **PHA** e **CH50%**.

A metodologia realizada na análise estrutural encontra-se resumida no diagrama abaixo (figura 2).

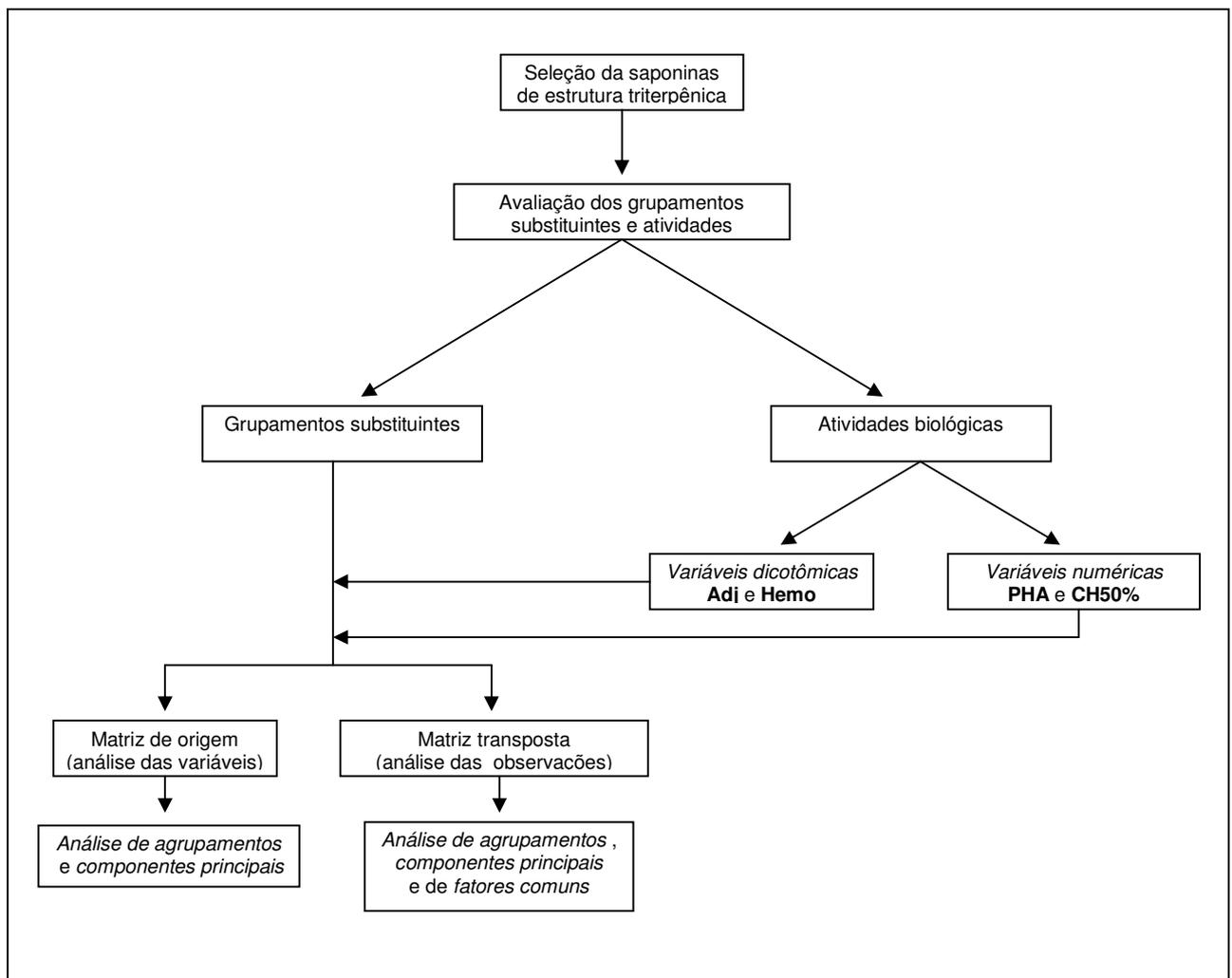


Figura 2. Diagrama da metodologia aplicada à análise estrutural.

3.1.5 Análise da capacidade preditiva do modelo proposto

Quatro saponinas codificadas (**SAP-1**, **SAP-2**, **SAP-3** e **SAP-4**), não utilizadas no estabelecimento do modelo, foram submetidas a uma análise estrutural, considerando-se os substituintes relacionados com as capacidades imunoadjuvante e hemolítica das saponinas, identificados, principalmente, via *análise de agrupamentos*, na presença de atividades biológicas representadas por variáveis numéricas, **PHA** e **CH50%**. A análise foi realizada por pessoa alheia ao trabalho, a qual teve acesso apenas aos dendrogramas contendo as atividades biológicas representadas por variáveis numéricas.

4 RESULTADOS E DISCUSSÃO

4 RESULTADOS E DISCUSSÃO

4.1 Análise estrutural

4.1.1 Seleção das saponinas

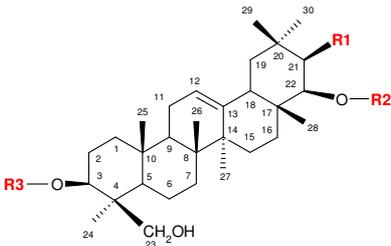
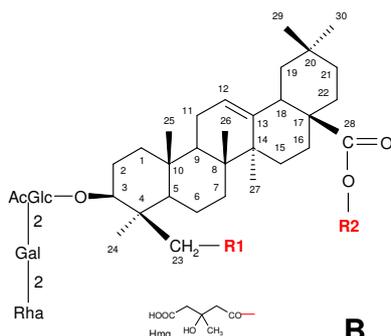
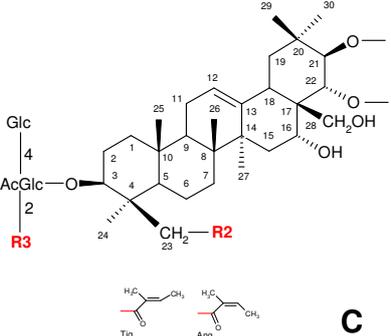
A seleção das saponinas incluídas no estudo foi realizada, principalmente, tendo por base o trabalho desenvolvido por ODA e col. (2000), onde foram avaliadas a atividade imunoadjuvante e hemolítica de 47 saponinas. A partir do referido estudo, foram selecionadas 20 saponinas (tabela 1) pertencentes aos grupos das sojasaponinas, dos lablabosídeos e das escinas derivadas de *G. max*, *D. lablab* e *A. hippocastanum*, respectivamente. O grupo de saponinas selecionado visou incluir aquelas com atividade hemolítica e imunoadjuvante, atividade hemolítica e sem atividade imunoadjuvante e vice-versa, e aquelas que não apresentam nenhuma dessas atividades. A restrição a um mesmo trabalho (o de ODA e col., 2000) se deveu à necessidade de comparar atividades imunoadjuvante e hemolítica determinadas pela mesma metodologia, evitando, dessa maneira, introduzir um viés no estudo.

Os compostos selecionados apresentam estruturas triterpênicas contendo variações estruturais usuais nesse tipo de saponina. Com isso, a diversidade de substituintes foi delimitada, de modo a evitar uma análise complexa de um número excessivo de variáveis.

O caráter restritivo imposto na escolha das saponinas faz com que o modelo preditivo proposto se limite a compostos com estrutura triterpênica, principalmente aos derivados do ácido oleanólico. Contudo, embora a amostra utilizada não seja representativa de todas as possíveis variações estruturais encontradas em saponinas, aquelas selecionadas incluem, sim, um subconjunto de grande importância, o que justifica o desenvolvimento de um modelo nos moldes propostos.

Resultados e Discussão

Tabela 1. Representação estrutural das saponinas triterpênicas derivadas de **(A)** *G. max*. **(B)** *D. lablab* e **(C)** *A. hippocastanum*.

Estrutura	Saponina (Abreviação)	R1	R2	R3
 <p style="text-align: center;">A</p>	Sojasapogenol A (SYOL - A)	OH	H	H
	Sojasapogenol B (SYOL - B)	H	H	H
	Sojasaponina I (SYA - 1)	H	H	- AcGlc – Gal – Rha
	Sojasaponina II (SYA - 2)	H	H	- AcGlc – Ara – Rha
	Sojasaponina III (SYA - 3)	H	H	- AcGlc – Gal
	Deidrosojasaponina I (DHSYA)	H	***	- AcGlc – Gal – Rha
	Sojasaponina A1 (SYA - A1)	OH	- Ara – Glc	- AcGlc – Gal – Glc
	Sojasaponina A2 (SYA - A2)	OH	- Ara – Glc	- AcGlc – Gal
 <p style="text-align: center;">B</p>	Lablabosideo A (LBL - A)	H	Glc	#
	Lablabosideo B (LBL - B)	OH	Glc	#
	Lablabosideo C (LBL - C)	OH	- Glc – Rha	#
	Lablabosideo D (LBL - D)	OH	- Glc – Hmg	#
	Lablabosideo E (LBL - E)	OH	- Glc – Rha – Rha	#
	Lablabosideo F (LBL - F)	H	- Glc – Rha – Rha	#
 <p style="text-align: center;">C</p>	Escina Ia (ESC - 1A)	Tig	OH	Glc
	Escina Ib (ESC - 1B)	Ang	OH	Glc
	Escina IIa (ESC - 2A)	Tig	OH	Xyl
	Escina IIb (ESC - 2B)	Ang	OH	Xyl
	Escina IIIa (ESC - 3A)	Tig	H	Glc
	Escina IIIb (ESC - 3B)	Ang	H	Glc

*** Deidrosojasaponina I apresenta uma carbolina no carbono 22 e portanto não apresenta substituinte em R2;

ausência de R3.

Monômeros presentes na cadeia de açúcares: Ácido Glicurônico (AcGlc), Arabinose (Ara), Glicose (Glc), Galactose (Gal), Ramnose (Rha) e Xilose (Xyl);

Grupos acila presentes: 3-hidroxi-3-metilglutaroil (Hmg), Angeloil (Ang) e Tigloil (Tig).

4.1.2 Avaliação dos grupamentos substituintes e atividades biológicas

Os grupamentos substituintes associados às saponinas selecionadas encontram-se discriminados na tabela 2. Neste primeiro momento, a cada grupo presente lhe foi conferido o mesmo grau de importância, visando eliminar qualquer tendência subjetiva.

Tabela 2. Grupamentos substituintes e abreviações das variáveis utilizadas na análise estrutural das saponinas triterpênicas.

Local de substituição	Grupamento substituinte	Abreviação
Carbono 3 (C-3)	Hidroxila (OH)	C3-OH
	Gal (1→2) AcGlc -	C3-a
	Rha(1→2)Gal (1→2) AcGlc -	C3-b
	Rha(1→2)Ara(1→2) AcGlc -	C3-c
	Gal(1→2)Rha(1→2) AcGlc -	C3-d
	Glc(1→2)Gal(1→2) AcGlc -	C3-e
	Glc(1→4) AcGlc (2←1)Xyl	C3-f
	Glc(1→4) AcGlc (2←1)Glc	C3-g
Carbono 4 (C-4)	Uma metila (CH ₃)	C4-met
	Duas metilas (CH ₃)	C4-2met
	Hidroxila (OH)	C4-OH
	CH ₂ OH	C4-metox
Carbono 16 (C-16)	Hidroxila (OH)	C16-OH
Carbono 17 (C-17)	Metila (CH ₃)	C17-met
	CH ₂ OH	C17-metox
	COO-ligado a um açúcar	C17-carbose
Carbono 21 (C-21)	Hidroxila (OH)	C21-OH
	Tigloil	C21-tig
	Angeloil	C21-ang
Carbono 22 (C-22)	Hidroxila (OH)	C22-OH
	Acetil (OCOCH ₃)	C22-carbmet
	Carbonila	C22-carb
	Glc(1→3) Ara -	C22-a
Carbono 28 (C-28)	Hidroxila (OH)	C28-OH
	Glc	C28-a
	Rha(1→4) Glc -	C28-b
	Hmg(1→6) Glc -	C28-c
	Rha(1→4)Rha(1→2) Glc -	C28-d

Ácido Glicurônico (AgGlc) , Arabose (Ara), Galactose (Glc) monômeros da porção osídica ligados a aglicona.

A análise dos substituintes foi realizada mediante codificação via variáveis dicotômicas, sendo que “1” significa presença de determinado grupamento e “0” a ausência do mesmo. Esse tipo de abordagem foi utilizado devido a sua compatibilidade com os modelos estatísticos posteriormente aplicados neste trabalho, permitindo uma diferenciação das *variáveis* (representadas pelos diferentes substituintes) de acordo com as *observações* consideradas (representadas pelas saponinas).

Além dos diferentes grupamentos, foram incluídas as atividades imunoadjuvante e hemolítica das saponinas disponíveis na literatura (ODA e col. 2000). Os valores relatados foram inseridos em termos quantitativos, mediante o título de **PHA**, no caso da atividade imunoadjuvante e mediante a **CH50%**, no caso da atividade hemolítica, e, portanto, representam variáveis numéricas. Já em termos qualitativos, as variáveis dicotômicas foram codificadas como **Adj** e **Hemo** e representam as atividades imunoadjuvante e hemolítica, respectivamente. As variáveis referentes às atividades biológicas foram tratadas desta forma, para que nos modelos estatísticos fosse possível, tanto a avaliação do comportamento das saponinas relacionado ao simples fato de apresentarem ou não atividade imunoadjuvante e hemolítica, como também a possibilidade de agrupar os compostos e as demais variáveis em virtude de sua potência imunoadjuvante e hemolítica.

Após a fixação das variáveis e das observações, procedeu-se à codificação dos grupamentos substituintes e das atividades biológicas (tabela 3). Essa base de dados foi utilizada para alimentar as análises estatísticas e, portanto, serve de pilar para o desenvolvimento do modelo preditivo proposto neste trabalho.

O **PHA** e **CH50%** representam variáveis numéricas e, portanto, possuem um comportamento diferenciado em relação às demais. Assim, ao analisarmos essas variáveis expressas pelas magnitudes originalmente utilizadas por ODA e col. (2000) em relação às outras que são codificadas, observou-se uma distorção no sistema levando a resultados irrealistas. Para contornar o problema, os dados referentes às variáveis numéricas foram transformados, criando-se para isso uma escala numérica com a sensibilidade necessária para refletir a potência imunoadjuvante e hemolítica das saponinas sem levar a distorções.

Tabela 3. Codificação das variáveis referentes aos substituintes e às atividades imunoadjuvante e hemolítica das saponinas triterpênicas provenientes de *A. hippocastanum*, *D. lablab* e *G. max*.

Saponinas	C3-OH	C3-a	C3-b	C3-c	C3-d	C3-e	C3-f	C3-g	C4-met	C4-2met	C4-metox	C16-OH	C17-met	C17-metox	C17-carbose	C21-OH	C21-tig
SYOL - A	1	0	0	0	0	0	0	0	1	0	1	0	1	0	0	1	0
SYOL - B	1	0	0	0	0	0	0	0	1	0	1	0	1	0	0	0	0
SYA - 1	0	0	1	0	0	0	0	0	1	0	1	0	1	0	0	0	0
SYA - 2	0	0	0	1	0	0	0	0	1	0	1	0	1	0	0	0	0
SYA - 3	0	1	0	0	0	0	0	0	1	0	1	0	1	0	0	0	0
DHSYA	0	0	0	0	1	0	0	0	1	0	1	0	1	0	0	0	0
SYA - A1	0	0	0	0	0	1	0	0	1	0	1	0	1	0	0	1	0
SYA - A2	0	1	0	0	0	0	0	0	1	0	1	0	1	0	0	1	0
LBL - A	0	0	1	0	0	0	0	0	0	1	0	0	0	0	1	0	0
LBL - B	0	0	1	0	0	0	0	0	1	0	1	0	0	0	1	0	0
LBL - C	0	0	1	0	0	0	0	0	1	0	1	0	0	0	1	0	0
LBL - D	0	0	1	0	0	0	0	0	1	0	1	0	0	0	1	0	0
LBL - E	0	0	1	0	0	0	0	0	1	0	1	0	0	0	1	0	0
LBL - F	0	0	1	0	0	0	0	0	0	1	0	0	0	0	1	0	0
ESC - 1A	0	0	0	0	0	0	0	1	1	0	1	1	0	1	0	0	1
ESC - 1B	0	0	0	0	0	0	0	1	1	0	1	1	0	1	0	0	0
ESC - 2A	0	0	0	0	0	0	1	0	1	0	1	1	0	1	0	0	1
ESC - 2B	0	0	0	0	0	0	1	0	1	0	1	1	0	1	0	0	0
ESC - 3A	0	0	0	0	0	0	0	1	0	1	0	1	0	1	0	0	1
ESC - 3B	0	0	0	0	0	0	0	1	0	1	0	1	0	1	0	0	0

Onde, 1 significa a **presença** de determinado grupamento e 0 significa **ausência**.

Tabela 3. Codificação das variáveis referentes aos substituintes e às atividades imunoadjuvante e hemolítica das saponinas triterpênicas provenientes de *A. hippocastanum*, *D. lablab* e *G. max.* (Cont.).

Saponinas	C21-ang	C22-OH	C22-carbmet	C22-carb	C22-a	C28-OH	C28-a	C28-b	C28-c	C28-d	PHA *	CH50% **	Adj [‡]	Hemo [°]
SYOL - A	0	1	0	0	0	0	0	0	0	0	15 (0,5)	≥ 600 (0,5)	0	0
SYOL - B	0	1	0	0	0	0	0	0	0	0	15 (0,5)	≥ 600 (0,5)	0	0
SYA - 1	0	1	0	0	0	0	0	0	0	0	210 (2)	≥ 600 (0,5)	1	0
SYA - 2	0	1	0	0	0	0	0	0	0	0	314 (3)	≥ 600 (0,5)	1	0
SYA - 3	0	1	0	0	0	0	0	0	0	0	240 (2)	≥ 600 (0,5)	1	0
DHSYA	0	0	0	1	0	0	0	0	0	0	240 (2)	≥ 600 (0,5)	1	0
SYA - A1	0	0	0	0	1	0	0	0	0	0	540 (4,5)	≥ 600 (0,5)	1	0
SYA - A2	0	0	0	0	1	0	0	0	0	0	130 (1,5)	≥ 600 (0,5)	1	0
LBL - A	0	0	0	0	0	0	1	0	0	0	54 (1)	≥ 600 (0,5)	1	0
LBL - B	0	0	0	0	0	0	1	0	0	0	31 (0,5)	≥ 600 (0,5)	1	0
LBL - C	0	0	0	0	0	0	0	1	0	0	53 (1)	≥ 600 (0,5)	1	0
LBL - D	0	0	0	0	0	0	0	0	1	0	90 (1)	≥ 600 (0,5)	1	0
LBL - E	0	0	0	0	0	0	0	0	0	1	41 (1)	≥ 600 (0,5)	1	0
LBL - F	0	0	0	0	0	0	0	0	0	1	470 (4)	≥ 600 (0,5)	1	0
ESC - 1A	0	0	1	0	0	1	0	0	0	0	42 (0,5)	128 (1)	1	1
ESC - 1B	1	0	1	0	0	1	0	0	0	0	16 (0,5)	57 (2)	0	1
ESC - 2A	0	0	1	0	0	1	0	0	0	0	21 (0,5)	28 (2)	0	1
ESC - 2B	1	0	1	0	0	1	0	0	0	0	21 (0,5)	36 (2)	0	1
ESC - 3A	0	0	1	0	0	1	0	0	0	0	21 (0,5)	22 (2)	0	1
ESC - 3B	1	0	1	0	0	1	0	0	0	0	21 (0,5)	28 (2)	0	1

Onde, **1** significa a **presença** de determinado grupamento e **0** significa **ausência**.

* Título de PHA obtido mediante teste de Hemaglutinação Passiva;

** Concentração hemolítica 50% (CH50%) em µg/ml;

‡ Atividade imunoadjuvante considerada presente quando o título de PHA for significativamente maior que o branco;

° Atividade hemolítica considerada presente quando a CH50% for menor que 600 µg/ml.

() Resultados das variáveis PHA e CH50% em escala.

4.1.3 Análise de agrupamentos

A *análise de agrupamentos* faz parte das *técnicas multivariadas* e tem como finalidade agregar observações com base nas características que elas possuem. Os *agrupamentos* resultantes devem, por definição, exibir elevada homogeneidade interna (dentro dos agrupamentos) e elevada heterogeneidade externa (entre agrupamentos). Esta técnica, entre uma gama de aplicações, pode realizar a redução dos dados de forma objetiva mediante sistematização das informações de uma população inteira, ou de uma amostra, gerando informações sobre subgrupos específicos e menores. O objetivo primário da análise de agrupamentos é dividir um conjunto de observações em dois ou mais grupos com base na similaridade das observações, em relação a um conjunto de características especificadas (HAIR e col., 2005).

Inserido neste contexto, a análise de agrupamentos foi realizada com o objetivo de resumir os dados obtidos referentes aos diferentes grupamentos substituintes e atividades biológicas (tabela 3), de modo a simplificar a análise dos resultados. Como esta técnica acaba criando diferentes subgrupos com características similares (agrupamentos), facilita a avaliação das relações estruturais das saponinas, envolvidas com suas atividades imunoadjuvante e hemolítica.

Na tentativa de ajustar o modelo proposto, analisou-se tanto a matriz de origem (conforme representado na tabela 3) como da matriz transposta, ou seja, as linhas que eram representadas pelas diferentes saponinas são transformadas em colunas e, portanto, passam a ser as **variáveis da matriz**, já com os diferentes substituintes ocorre o processo inverso, sendo que acabam representando as **observações da matriz**.

A figura 3 representa a análise comparativa entre variáveis obtidas da matriz de origem, e observações obtidas da matriz transposta na presença e ausência de atividades biológicas codificadas (**Adj** e **Hemo**). Mediante esse procedimento se fez possível analisar os diferentes agrupamentos formados quando as atividades biológicas codificadas são inseridas entre as variáveis de entrada.

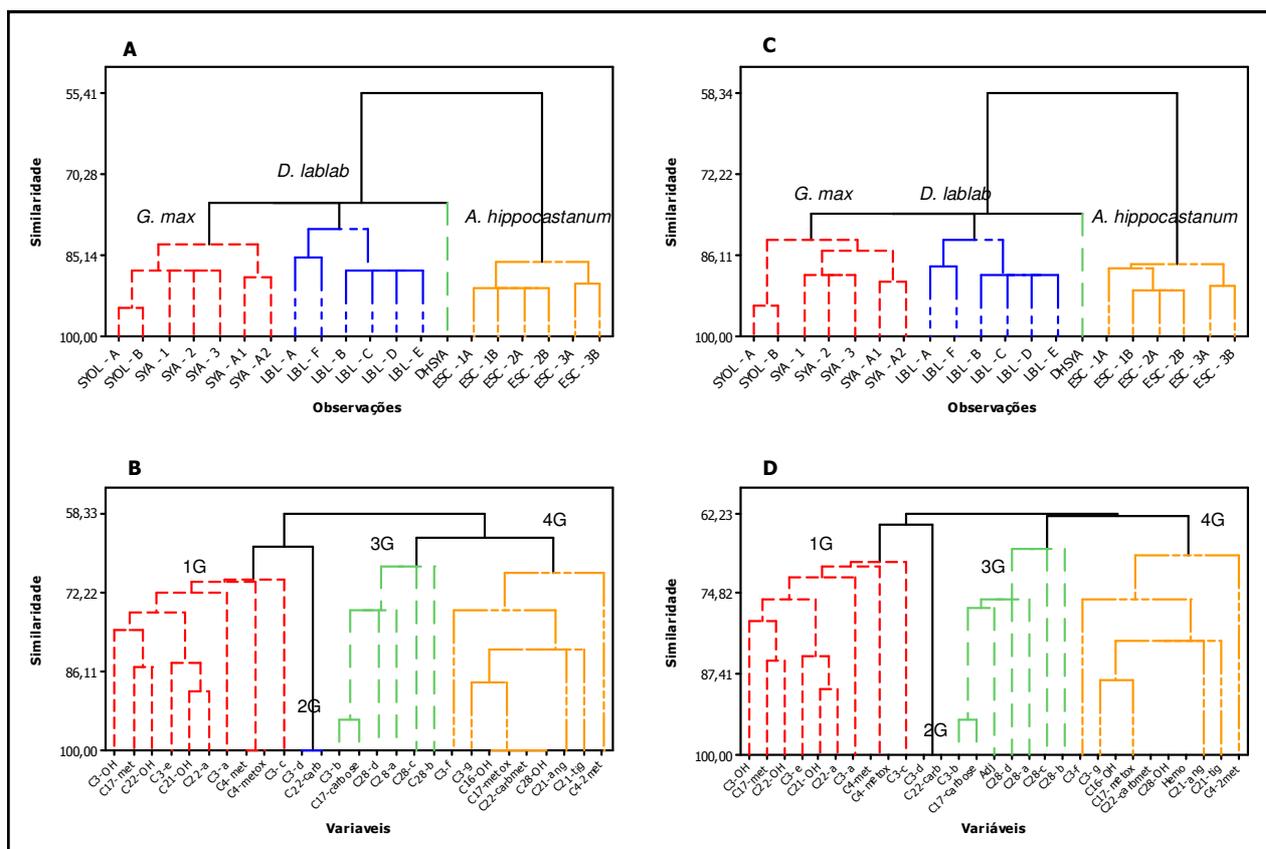


Figura 3. Dendrogramas das observações e variáveis na (A e B) ausência e (C e D) presença de atividades biológicas codificadas.

A análise da figura 3A revela a presença de quatro agrupamentos distintos no dendrograma que representa as observações na ausência de atividades biológicas. O primeiro grupamento é composto pelas saponinas de *G. max*, sendo que são encontradas três subpopulações nesse agrupamento, representadas pelas agliconas (**SYOL-A** e **SYOL-B**), pelas saponinas monodesmosídicas (**SYA-A**, **SYA-B** e **SYA-C**) e pelas saponinas bidesmosídicas (**SYA-A1** e **SYA-A2**), com níveis de similaridade superiores a 86%. O segundo agrupamento é composto pelas saponinas de *D. lablab*, sendo que são encontradas duas subpopulações representadas por **LBL-A** e **LBL-F**, que são as únicas no agrupamento a apresentarem duas metilas em C-4, enquanto a segunda subpopulação envolve os demais lablabosídeos que apresentam uma metila e um CH₂OH em C-4. O terceiro agrupamento apresenta apenas uma saponina (**DHSYA**), derivada de *G. max*, que possui características estruturais próprias, como a carbonila em C-22, que a diferencia das demais saponinas de soja, formando um

agrupamento separado. O quarto agrupamento é composto pelas saponinas de *A. hippocastanum*, onde são encontradas duas subpopulações representadas, uma pela **ESC-3A** e **ESC-3B** que apresentam duas metilas em C-4, e a outra pelas demais escinas que apresentam uma metila e um CH₂OH em C-4.

De um modo geral, os diferentes agrupamentos obtidos (figura 3A) indicam que as saponinas que compõem esses grupos apresentam, de fato, características estruturais semelhantes e são derivadas das mesmas espécies vegetais. Já as subpopulações dos lablabosídeos e escinas são definidas, dentro de um mesmo agrupamento, pelos substituintes em C-4. Porém, nas saponinas de soja essas subpopulações são determinadas basicamente pelo número de cadeias de açúcares ligados à aglicona, uma vez que todas apresentam o mesmo substituinte em C-4. Já a posição do ponto de bifurcação que liga uma subpopulação a outra é, geralmente, determinado pela semelhança entre os açúcares ligados a aglicona.

No momento em que as atividades são incluídas como variáveis dicotômicas (**Adj** e **Hemo**) na matriz, o dendrograma apresenta praticamente o mesmo perfil (figura 3C). Evidenciam-se, contudo, pequenas alterações com relação à similaridade, como no primeiro agrupamento, onde as agliconas de *G. max* (**SYOL-A** e **SYOL-B**), que não apresentam atividade imunoadjuvante, passam a ter uma menor similaridade em relação as saponinas monodesmosídicas e bidesmosídicas. As saponinas monodesmosídicas e bidesmosídicas de soja têm a sua similaridade aumentada, já que ambas são imunoadjuvantes e não-hemolíticas. No segundo agrupamento, ocorre um aumento de similaridade entre os membros devido ao fato de todos eles apresentarem atividade imunoadjuvante, mas não atividade hemolítica. O quarto agrupamento sofreu leves modificações, sendo que a **ESC-1A** passa a formar uma subpopulação distinta, por ser a única escina a apresentar tanto atividade imunoadjuvante, como hemolítica, já que as demais escinas apresentam apenas atividade hemolítica.

Considerando a análise das variáveis na ausência das atividades biológicas (figura 3B), podem-se observar no dendrograma quatro agrupamentos distintos, que denominamos 1G, 2G, 3G e 4G. O primeiro (1G) é composto basicamente por variáveis comuns às saponinas de *G. max*, excetuando a **DHSYA**. As variáveis **C17-met** e **C22-OH** apresentam uma similaridade de aproximadamente 86%, sendo ambas variáveis

comuns nas saponinas monodesmosídicas de soja selecionadas. Já **C21-OH** e **C22-a** apresentam uma similaridade de aproximadamente 90%, sendo ambas comuns nas saponinas bidesmosídicas de soja **SYA-A1** e **SYA-A2**, além de estarem presentes na aglicona **SYOL-A**, da qual as saponinas bidesmosídicas anteriormente citadas são derivadas. Entre **C4-met** e **C4-metox** podemos observar uma similaridade de 100%, o que reflete a total correlação entre as variáveis, já que são dependentes uma em relação à outra. O segundo agrupamento (2G) é composto por **C3-d** e **C22-carb** que representam o açúcar em C-3 e a carbonila em C-22, que são variáveis exclusivas da **DHSYA** e possuem, portanto, uma similaridade de 100%. O terceiro agrupamento (3G) é composto basicamente por variáveis associadas aos lablabosídeos, já que essas saponinas possuem características estruturais semelhantes. As variáveis **C3-b** e **C17-carbose** apresentam uma similaridade de aproximadamente 90%, uma vez que ambas estão presentes em todas as saponinas derivadas de *D. lablab*. Já as demais variáveis que fazem parte do agrupamento, representam açúcares que são, geralmente, diferentes entre os membros do grupo e, portanto, apresentam uma menor e diferenciada similaridade. O quarto agrupamento (4G) é composto, basicamente, por variáveis associadas às escinas, sendo que **C16-OH**, **C17-metox**, **C22-carbmet** e **C28-OH** apresentam uma similaridade de 100%, uma vez que estão presentes em todos os membros desse grupo. O trissacarídeo **C3-g**, que é um açúcar muito comum às saponinas de *A. hippocastanum*, a exceção de **ESC-2A** e **2B**, apresenta um alto nível de similaridade em relação às variáveis exclusivas ao grupo das escinas.

A introdução das atividades biológicas na matriz (figura 3D) pouco altera a estrutura geral do dendrograma que se mantém constante, apresentando apenas pequenas alterações na similaridade dos agrupamentos. Porém, pode-se perceber que a variável **Adj**, que representam a atividade imunoadjuvante, agrupa-se com as características relacionadas ao lablabosídeos, cujos membros apresentam atividade imunoadjuvante, característica que não se repete nos demais agrupamentos. Além disso, a variável **Adj** mantém uma relação próxima com **C3-b** e **C17-carbose** que são substituintes comuns às saponinas derivadas de *D. lablab*. Já a variável **Hemo**, que representa a atividade hemolítica, é agrupada intimamente com características comuns

a todas as escinas (**C16-OH**, **C17-metox**, **C22-carbmet** e **C28-OH**), sendo estas as únicas saponinas dentro da população em estudo a apresentarem atividade hemolítica.

Desta forma, essa avaliação preliminar indica que o método é sensível a introdução de variáveis que representam as atividades biológicas, gerando apenas pequenas modificações nos dendrogramas que descrevem às observadas nas variáveis de entrada. Além disso, pode-se visualizar que os membros dos agrupamentos apresentam características estruturais comuns que justificam as associações formadas.

A avaliação do comportamento dos agrupamentos na presença das variáveis numéricas **PHA** e **CH50%**, que representam as potências biológicas das saponinas, foi realizada mediante análises das variáveis e observações, sendo os resultados representados na figura 4.

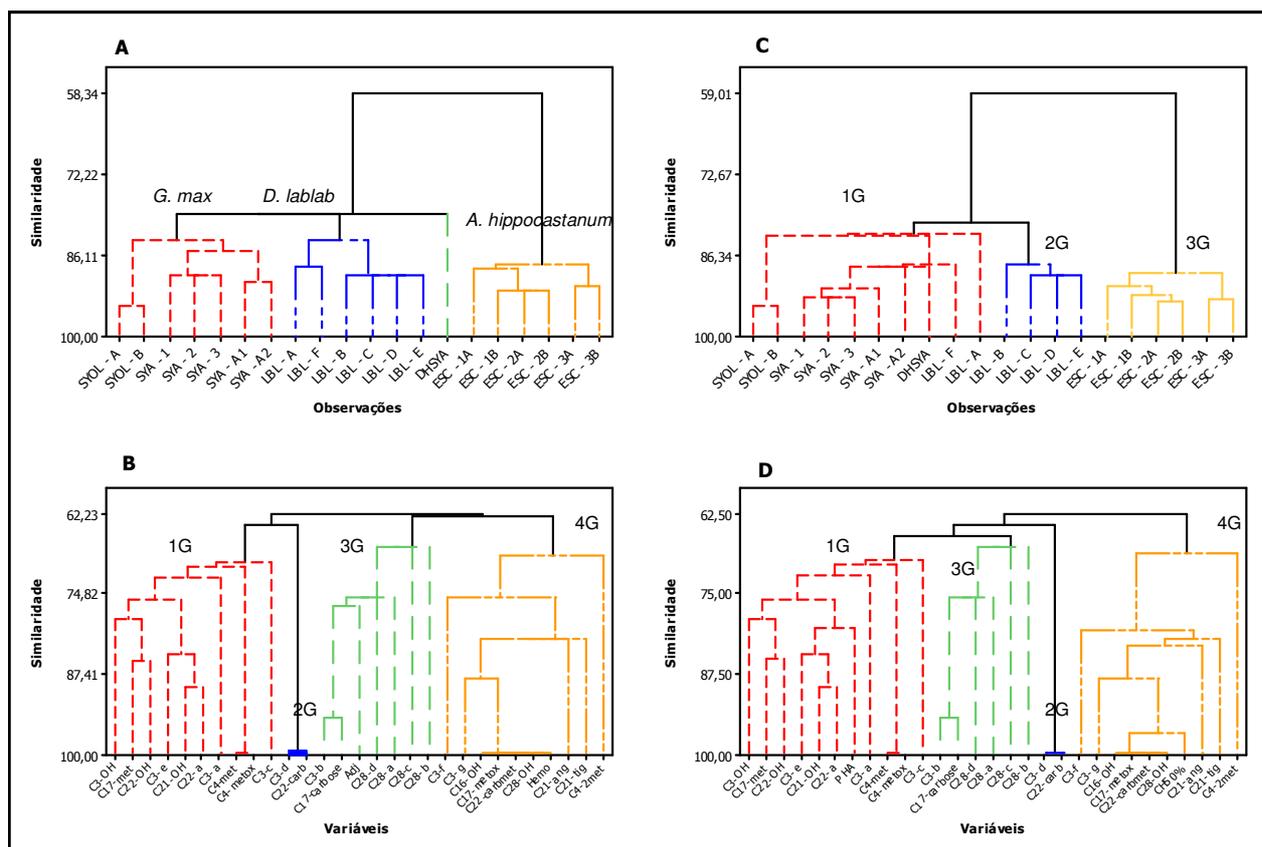


Figura 4. Dendrogramas das observações e das variáveis na (A e B) presença de atividades biológicas codificadas e na (C e D) presença de atividades biológicas representadas por variáveis numéricas.

Avaliando-se o comportamento do dendrograma na análise das observações após a introdução das variáveis **PHA** e **CH50%** (figura 4C), pode-se visualizar uma reorganização dos agrupamentos, se comparado ao dendrograma da figura 4A. A saponina **SYA-A1** aumenta sua similaridade em relação às saponinas monodesmosídicas de soja (1G), refletindo o elevado potencial imunoadjuvante comum a todas elas, contudo em relação à **SYA-A2**, sua similaridade diminui, visto que esta saponina possui um baixo potencial imunoadjuvante. A **DHSYA** devido ao seu **PHA** semelhante às saponinas monodesmosídicas de soja passa a fazer parte do primeiro agrupamento. Já o **LBL-F** em virtude de seu alto potencial imunoadjuvante, semelhante à **SYA-A1**, também passa a fazer parte do primeiro agrupamento. O **LBL-A**, que antes aparecia fortemente relacionado com **LBL-F** (figura 4A), refletindo o fato de ambas apresentarem duas metilas em C-4, passa a apresentar uma baixa similaridade em relação ao **LBL-F** e demais membros do agrupamento (figura 4C), refletindo assim sua menor atividade imunoadjuvante dentro da subpopulação. O segundo agrupamento (2G) é formado pelas demais saponinas de *D. lablab*, sendo que diminui o nível de similaridade do **LBL-B** em relação ao **LBL-C**, **LBL-D** e **LBL-E**, já que **LBL-B** apresenta a menor atividade imunoadjuvante se comparado aos demais. O terceiro agrupamento (3G) é formado pelas escinas. Neste pode-se observar que ocorre uma segregação na subpopulação anteriormente formada por **ESC-1B**, **ESC-2A** e **ESC-2B** (figura 4A), na qual a **ESC-1B** tem a sua similaridade diminuída em relação às demais (figura 4C), resultado da sua menor atividade hemolítica em relação a **ESC-2A** e **ESC-2B**.

As modificações que ocorrem no dendrograma das observações após a introdução das atividades biológicas, representadas pelas variáveis numéricas **PHA** e **CH50%** (figura 4C), demonstram a sensibilidade do método à introdução destas a matriz, assim como a sua capacidade discriminante em relação a potência imunoadjuvante e hemolítica das saponinas. O primeiro agrupamento (1G) é formado basicamente por saponinas que apresentam alta capacidade imunoadjuvante, a exceção da **SYOL-A** e **SYOL-B**, que são agliconas e apresentam baixa capacidade imunoadjuvante e hemolítica. O segundo agrupamento (2G) é composto basicamente por estruturas que apresentam um potencial imunoadjuvante moderado. Já o terceiro

(3G) é formado por saponinas que apresentam elevado potencial hemolítico e baixo potencial imunoadjuvante.

Em relação às variáveis, pode-se visualizar mediante análise do dendrograma (figura 4D), que ocorrem pequenas modificações nos níveis de similaridade sem alterações na estrutura geral, se comparado ao dendrograma da figura 4B. A variável **PHA** se agrupa com as características comuns às saponinas derivadas de soja, refletindo a maior capacidade imunoadjuvante dessas saponinas se comparadas às demais. Contudo, a variável **Adj** se agrupa com características dos lablabosídeos, refletindo o fato de todos apresentarem atividade imunoadjuvante, o que não acontece com as estruturas derivadas de soja devido à inclusão das agliconas na análise (**SYOL-A** e **SYOL-B**) que não apresentam atividade imunoadjuvante.

Na análise dos diferentes agrupamentos formados (figura 4D), pode-se observar que a variável **PHA** está relacionada principalmente com **C22-a**, **C21-OH** e **C3-e**, visto que estas características são encontradas na **SYA-A1**, que apresenta a maior capacidade imunoadjuvante na população analisada. Já a **CH50%** possui uma relação muito íntima com as variáveis comuns às escinas, em especial, **C16-OH**, **C17-metox**, **C22-carbmet** e **C28-OH**.

A definição final do número de grupos a ser formado é uma questão delicada na *análise de agrupamentos*. O critério adotado no presente trabalho corresponde à denominada *regra de parada*, a qual simplesmente considera a *medida de similaridade* em cada passo sucessivo dentro do *procedimento hierárquico aglomerativo*. Segundo esse critério, a solução de agrupamentos é definida quando a medida de similaridade excede a um valor especificado ou quando os valores sucessivos entre etapas possuem um salto repentino. Além da fundamentação teórica envolvida com a regra de parada, existe o julgamento crítico e o senso comum que devem ser analisados no momento de se determinar o número final de agrupamentos (HAIR e col., 2005).

Neste contexto, extraíram-se apenas três agrupamentos na análise das observações na presença de atividades biológicas representadas pelas variáveis numéricas (figura 4C), visto que essa solução permite explicar melhor os resultados observados nas variáveis de entrada, se comparada à de quatro agrupamentos utilizada nas demais análises. Dessa forma, se chega à formação de agrupamentos que

refletem de forma coerente o potencial imunoadjuvante e hemolítico das saponinas analisadas.

Os dendrogramas analisados revelam que os diferentes agrupamentos formados entre as saponinas refletem os grupos formados entre as variáveis. Portanto, a técnica mostrou-se sensível as modificações nas variáveis de entrada sendo capaz de facilitar a análise das relações estruturais envolvidas com as atividades biológicas, inclusive com as potências biológicas, tornando-se um método factível de ser utilizado em estudos envolvendo a busca de padrões estruturais que determinam as atividades biológicas de saponinas, podendo ser estendido a outros estudos relacionados com avaliações estruturais.

4.1.4 Análise de fatores

A *análise de fatores* possui dois objetivos básicos, a saber, a identificação da estrutura por meio de um conjunto de dados e a própria redução de dados. Além disso, fornece uma visão muito direta das inter-relações entre variáveis ou observações, bem como apoio empírico para abordar questões conceituais relativas à estrutura latente dos dados (HAIR e col., 2005).

Esta técnica estatística multivariada possui dois modelos básicos para obtermos soluções fatoriais, conhecidos como *análise de componentes principais* e *análise de fatores comuns*.

4.1.4.1 Análise de componentes principais

Um dos critérios primordiais nesta *técnica multivariada* é a determinação do número de *componentes* a serem extraídos. Para isso existem diversas técnicas que visam determinar a combinação de *componentes* capaz de explicar o máximo da variância intrínseca à população sob estudo. Porém, esses métodos nos fornecem informações puramente técnicas, que, dependendo da situação, podem não refletir a

melhor representação dos dados (HAIR e col., 2005). Portanto, o número de *componentes* a serem extraídos deve ser determinado mediante associação das informações técnicas e do julgamento crítico, assim como o número de *agrupamentos* na *análise de agrupamentos*. Neste contexto, o número de componentes a serem extraídos foi baseado no *critério da raiz latente*, no qual, para fins de interpretação, qualquer componente individual deve explicar a variância de pelo menos uma variável. Cada variável contribui com um valor unitário do autovalor total. Logo, são considerados significantes apenas os componentes que têm raízes latentes ou autovalores maiores que a unidade (HAIR e col., 2005).

Na análise de componentes principais, seguiu-se a mesma sistemática na abordagem das variáveis de entrada utilizadas na análise de agrupamentos, ou seja, foram realizadas análises da matriz de origem e da matriz transposta, meio pelo qual foram avaliadas as intercorrelações entre as variáveis e as observações, respectivamente.

Para avaliar comparativamente o comportamento das variáveis e observações perante a introdução das variáveis dicotômicas que representam as atividades biológicas (**Adj** e **Hemo**), optou-se pela análise comparativa onde, primeiramente, foram considerados apenas os grupamentos substituintes. Os resultados referentes ao critério da raiz latente podem ser visualizados na figura 5.

A análise dos gráficos de autovalor revelou que, no caso das observações (figura 5A e 5C), os autovalores são superiores à unidade até o quarto componente, tanto na ausência como na presença de atividades biológicas codificadas. Com isso o número de componentes a ser extraído foi fixado em quatro. Já no caso das variáveis (figura 5B e 5D), a partir do componente número 11 os autovalores foram inferiores à unidade, tanto na ausência como na presença de atividades biológicas codificadas, levando, conseqüentemente, à fixação de 11 componentes a serem extraídos.

Quando um grande conjunto de variáveis é transformado em componentes, o método extrai primeiro as combinações de variáveis que explicam o maior montante da variância e depois segue para combinações que explicam montantes cada vez menores da variância. Conseqüentemente, os dois primeiros componentes são os de maior importância, por expressarem as causas de maior variabilidade entre as variáveis, ou

seja, a maior *variância comum*. Já os demais componentes explicam uma proporção cada vez menor da *variância comum*, sendo que os últimos expressam apenas pequenas proporções de *variância comum* e grandes proporções de *variância específica*, também conhecida como *variância única*, que está relacionada com uma variável específica (HAIR e col., 2005).

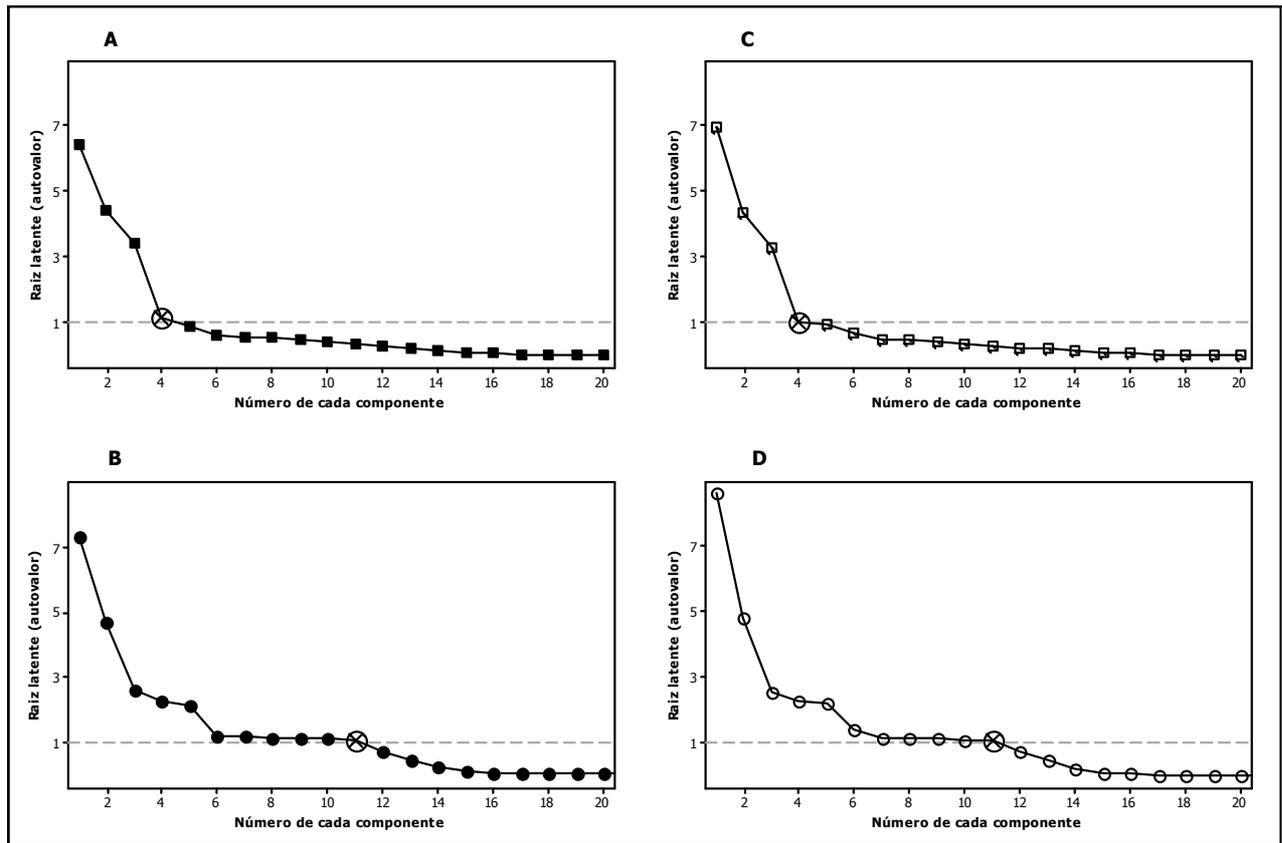


Figura 5. Gráficos de autovalor das observações e variáveis na (A e B) ausência e na (C e D) presença de atividades biológicas codificadas.

Os gráficos referentes à análise de componentes principais, das observações e variáveis na ausência e presença de atividades biológicas codificadas tendo como critério a raiz latente, podem ser visualizados na figura 6.

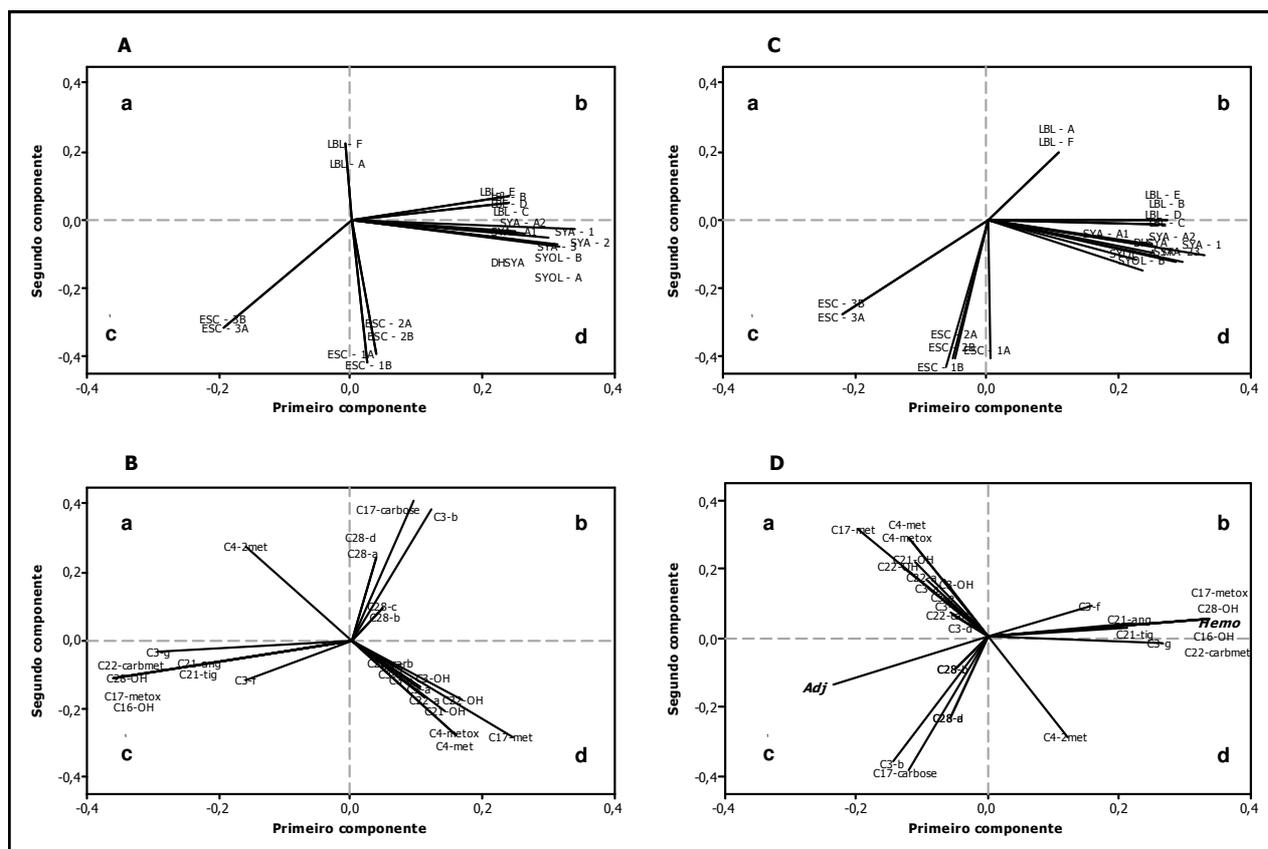


Figura 6. Representação gráfica da análise de componentes principais das observações e variáveis na (A e B) ausência e (C e D) na presença de atividades biológicas codificadas.

Pela análise do gráfico que representa as observações na ausência de atividades biológicas (figura 6A), observa-se no quadrante **c** uma forte correlação entre **ESC-3A** e a **ESC-3B**, refletindo o fato de ambas apresentarem duas metilas em C-4, característica que não se repete nas demais escinas, as quais apresentam uma metila e um CH₂OH em C-4 e que se traduz na correlação observada no quadrante **d**. Já no quadrante **a**, podemos observar uma forte correlação entre **LBL-A** e **LBL-F** que estão segregados dos demais lablabosídeos. Essas saponinas apresentam duas metilas em C-4, diferentemente das demais saponinas derivadas de *D. lablab*, **LBL-B**, **LBL-C**, **LBL-D** e **LBL-E**, que apresentam um CH₂OH e uma metila em C-4, o que justifica a correlação observada no quadrante **b**. Em relação ao quadrante **d**, existe uma forte correlação entre as saponinas derivadas de *G. max*, visto que todas apresentam um CH₂OH e uma metila em C-4.

A técnica de componentes principais aplicada às observações na ausência de atividades biológicas confirma os resultados obtidos na análise de agrupamentos, já que as correlações distintas observadas nos diferentes quadrantes são determinadas pelos substituintes em C-4, assim como, as diferentes subpopulações no mesmo grupo, na análise de agrupamentos.

Com a adição das variáveis **Adj** e **Hemo** às variáveis de entrada (figura 6C), ocorrem pequenas modificações, como a segregação da **ESC-1A** no quadrante **d**, já que esta é a única das escinas a apresentar atividade imunoadjuvante e hemolítica, visto que as demais apresentam apenas atividade hemolítica. Os vetores referentes ao **LBL-F** e **LBL-A** sofrem um pequeno desvio no sentido dos demais lablabosídeos no quadrante **b**, visto que todas as saponinas derivadas de *D. lablab* são imunoadjuvantes e não-hemolíticas. Esses desvios vetoriais, quando ocorrem no sentido dos demais membros da família, refletem os aumentos de similaridades entre uma subpopulação e outra, observados na análise de agrupamentos quando as atividades biológicas codificadas são adicionadas, caso o desvio ocorra no sentido inverso, significa que a similaridade diminui. Seguindo essa linha de pensamento, deveriam ser observados desvios nos vetores que representam as saponinas derivadas de soja, que refletissem as modificações de similaridade observadas na análise de agrupamentos. Porém, a sobreposição dos vetores prejudica visualização dos desvios. A pequena aproximação das saponinas derivadas de *G. max* em relação aos lablabosídeos reflete o pequeno aumento de similaridade observado pela análise de agrupamentos.

Pela análise das variáveis na ausência de atividades biológicas (figura 6B), observa-se uma distribuição de vetores que engloba os quatro quadrantes. No quadrante **b**, podemos observar uma forte correlação entre as variáveis comuns aos lablabosídeos. Já no quadrante **d**, existe uma forte correlação entre as variáveis comuns as saponinas derivadas de *G. max*. No quadrante **c**, podemos observar uma forte correlação entre as variáveis comuns às escinas. O quadrante **a** é ocupado pela variável **C4-2met** que se dilui ao longo das estruturas analisadas, não apresentando qualquer correlação direta com as demais variáveis. As correlações observadas confirmam os resultados obtidos pela análise de agrupamentos, onde os diferentes quadrantes ocupados refletem os agrupamentos formados e os desvios de um vetor em

relação ao outro no mesmo quadrante, geralmente, refletem as diferentes subpopulações observadas na análise de agrupamentos, conjuntamente com o nível de similaridades de uma variável em relação à outra.

Ao adicionarem-se as atividades biológicas codificadas (**Adj** e **Hemo**) às variáveis de entrada (figura 6D), ocorre uma inversão nos quadrantes, em comparação ao gráfico da figura 6B, devido a reorganização entre os componentes, principalmente em relação ao primeiro e segundo, já que explicam o maior montante da variância comum e, portanto, são os de maior importância. As variáveis referentes às escinas, que anteriormente localizavam-se no quadrante **c**, passam ao quadrante **b**, que é oposto ao anterior, visto que, após a introdução das variáveis **Hemo** e **Adj**, o primeiro componente se relaciona principalmente às características comuns às saponinas de *A. hippocastanum* que apresentam atividade hemolítica. Este fato é comprovado pelo sentido inverso dos vetores referentes aos substituintes comumente encontrados nos lablabosídeos e saponinas de soja, se comparadas às escinas, uma vez que as saponinas de *G. max* e *D. lablab* não são hemolíticas. O sentido oposto dos vetores que representam as variáveis **Adj** e **Hemo**, reforça tal suposição, revelando a existência de uma relação de oposição entre as variáveis e, conseqüentemente, entre as atividades imunoadjuvante e hemolítica. A variável **Adj** apresenta uma maior correlação com as variáveis comuns aos lablabosídeos, fato evidenciado pela maior proximidade dos vetores, dado que todos os membros dessa família apresentam atividade imunoadjuvante, característica que não se repete nos demais membros da população em análise. Já o segundo componente passa se relacionar principalmente com as variáveis comuns as saponinas de *G. max* (figura 6D), sendo que anteriormente estava relacionado a características comuns aos lablabosídeos (figura.6B)

Essas modificações que ocorrem após a introdução das atividades biológicas codificadas (**Adj** e **Hemo**) comprovam a sensibilidade do método e sua capacidade discriminativa quando são incluídas as variáveis de entrada, variáveis que representam as atividades biológicas.

Para a avaliação a capacidade discriminativa do método, em relação às potências biológicas das saponinas, as variáveis **PHA** e **CH50%** foram introduzidas como variáveis de entrada, procedendo-se as análises das observações e das

variáveis. O número de componentes a serem extraídos fundamenta-se na análise dos gráficos de autovalores (figura 7).

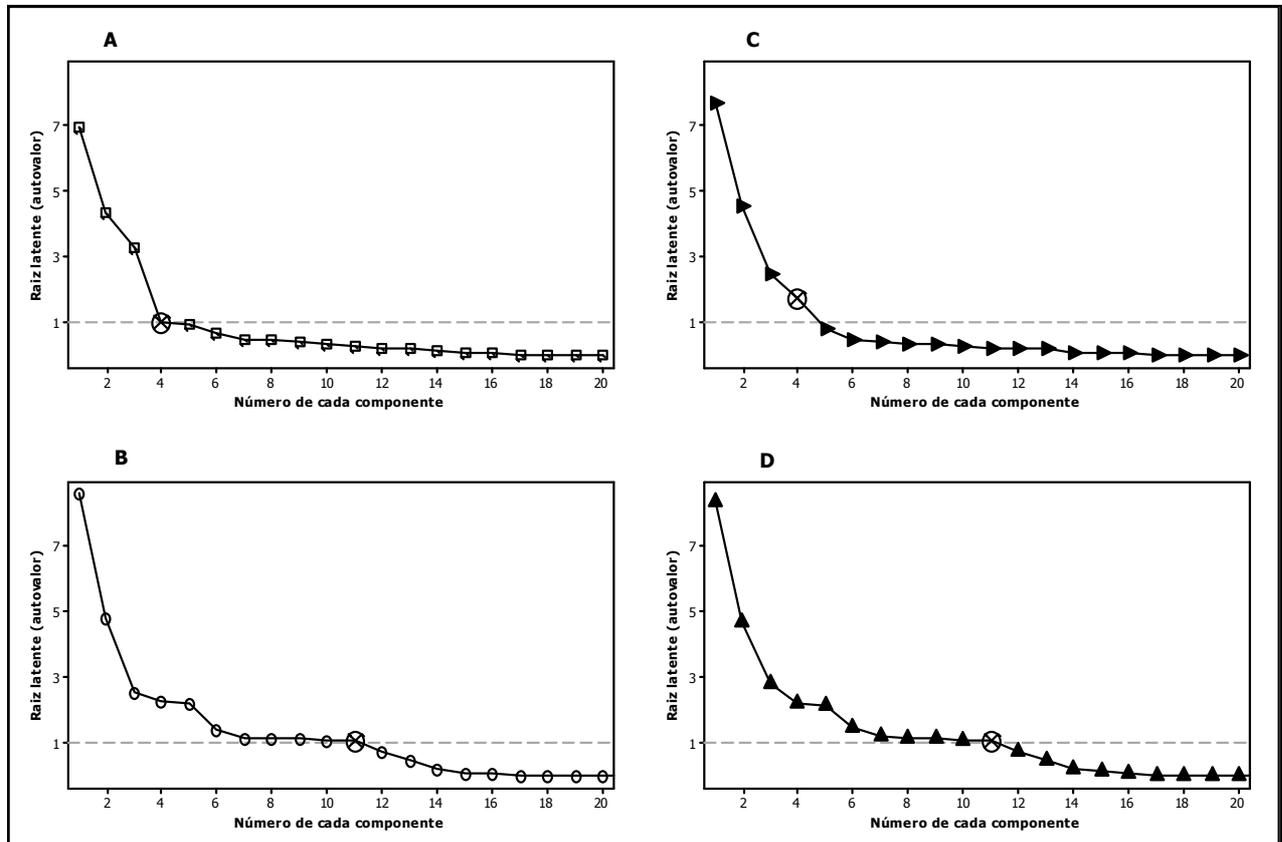


Figura 7. Gráficos de autovalor das observações e variáveis na (A e B) presença de atividades biológicas codificadas e na (C e D) presença de atividades biológicas representadas por variáveis numéricas.

Os gráficos de autovalor revelam que, após a introdução das variáveis numéricas, o número de componentes a serem extraídos manteve-se inalterado, tanto no caso das observações (quatro componentes) como no das variáveis (11 componentes), quando comparado aos gráficos da figura 5. Isso nos indica que a introdução das variáveis **PHA** e **CH50%** não aumenta exacerbadamente a variância entre as variáveis e que essas variáveis numéricas possuem o mesmo grau de importância se comparadas às variáveis dicotômicas que representam os diferentes substituintes, impedindo que ocorra a *saturação* do sistema de análise. Além disso, tal comportamento revela que a escala utilizada na adequação das variáveis numéricas é

ESC-3B e a outra subpopulação pelas demais saponinas derivadas de *A. hippocastanum*. Em virtude da grande sobreposição de vetores a visualização de modificações mais leves que possuem significado prático fica prejudicada.

Na análise das variáveis, após a introdução das variáveis numéricas **PHA** e **CH50%**, pode-se observar que a estrutura do gráfico (figura 8D) sofre apenas leves modificações quando comparado ao gráfico da figura 8B. Em comparação com a variável **Adj**, ocorre um desvio do vetor que representa a variável **PHA**, no sentido das características comuns às saponinas derivadas de *G. max*, as quais apresentam os maiores potenciais imunoadjuvantes. As características comuns às escinas, principalmente **C16-OH**, **C17-metox**, **C22-carbmet**, e **C28-OH**, se correlacionam fortemente com a variável **CH50%**, já que são as únicas saponinas dentro do conjunto analisado a apresentarem atividade hemolítica. Os módulos dos vetores das variáveis relacionadas a atividade hemolítica (**C16-OH**, **C17-metox**, **C22-carbmet**, e **C28-OH**) são elevados e no mesmo sentido do vetor da variável **CH50%**, indicando que estas variáveis podem influenciar a atividade e a capacidade hemolítica das saponinas analisadas, fato comprovado pela análise de agrupamentos. Além disso, a atividade imunoadjuvante pode ser influenciada pelas variáveis **C17-met**, **C3-b** e **C17-carbose**, que apresentam vetores de elevado módulo nos mesmos sentidos do vetor que representa a variável **PHA**.

A análise de agrupamentos revelou que a capacidade imunoadjuvante das saponinas pode estar relacionada à presença dos substituintes **C22-a**, **C21-OH** e **C3-e**. Já a análise de componentes principais foi capaz de revelar outros substituintes, **C17-met**, **C3-b** e **C17-carbose**, que também podem estar envolvidos com a capacidade imunoadjuvante das saponinas. Cruzando-se as informações obtidas, as variáveis **C22-a**, **C3-e**, **C3-b** e **C17-carbose** indicam que a capacidade imunoadjuvante das saponinas pode estar relacionada ao número de açúcares ligados a aglicona, sendo que as bidesmosídicas tendem a apresentar um maior potencial imunoadjuvante, quando comparadas às monodesmosídicas, indicando que a atividade imunoadjuvante pode ser influenciada pelas características anfipáticas das saponinas, conforme relatado anteriormente por ODA e col. (2003) para as saponinas de soja. Porém, comparando-se as saponinas monodesmosídicas de soja (**SYA-1**, **SYA-2** e **SYA-3**) com

as saponinas de *D. lablab*, que são bidesmosídicas, observa-se que as saponinas de soja, embora sejam monodesmosídicas, apresentam uma maior capacidade imunoadjuvante o que pode estar relacionado à presença de **C17-met**, presente nas saponinas de soja e ausente nos lablabosídeos. Portanto a capacidade imunoadjuvante das saponinas pode ser influenciada pela presença da hidroxila em C-21, da metila em C-17 e pela presença de dois açúcares ou cadeia de açúcares ligados à aglicona.

Os vetores que representam as variáveis **PHA** e **CH50%** possuem sentidos opostos, sendo que o ângulo formado entre eles é muito próximo a 180°, indicando uma forte relação de oposição entre as variáveis, ou seja, quanto maior o potencial imunoadjuvante de uma saponina menor o seu potencial hemolítico.

Existe uma relação de oposição entre os vetores tanto no caso das variáveis **Adj** e **Hemo** (figura 8B), como também entre as variáveis **PHA** e **CH50%** (figura 8D) indicando a existência de uma relação inversa entre as atividades imunoadjuvante e hemolítica para a população de saponinas avaliada. Mesmo no conjunto de saponinas analisadas existe uma exceção, a **ESC-1A**, que apresenta tanto atividade imunoadjuvante como hemolítica. Porém é inegável que existe uma relação de oposição entre as atividades analisadas, ou seja, saponinas que apresentam atividade imunoadjuvante tem uma menor tendência a apresentar atividade hemolítica. Segundo dados da literatura, não existe qualquer relação entre a atividade imunoadjuvante e hemolítica, conclusão oriunda principalmente da análise de resultados obtidos com saponinas de estrutura peculiar como, por exemplo, a **QS-21**, que apresentam elevada atividade imunoadjuvante e hemolítica (ODA e col.,2000).

O grupamento acila presente na estrutura de muitas saponinas pode estar relacionado à propriedade hemolítica (MATSUDA e col., 1997). Na população de saponinas analisada, todas as escinas apresentam em C-21 um grupamento acila (**C21-tig** ou **C21-ang**) e também apresentam atividade hemolítica. No entanto, o **LBL-D** apresenta o grupamento acila (3-hidroxi-3metilglutaroil) ligado ao açúcar em C-28, mas não apresenta atividade hemolítica, o que levou ODA e col. (2000) a concluírem que a atividade hemolítica não esta relacionada à presença de um grupamento acila. Entretanto, pode-se observar mediante análise da figura 8D, que as variáveis **C21-tig** e **C21-ang**, apresentam vetores de módulos discretos (a presença do grupamento acila

nas saponinas foi avaliada por meio das duas variáveis em separado), mas no mesmo sentido do vetor que representa a variável **CH50%**, o que indica uma forte correlação entre as variáveis, revelando que a presença do grupamento acila pode estar intimamente ligada à capacidade hemolítica das saponinas na população analisada. Além disso, o **LBL-D** apresenta o grupamento acila ligado a um açúcar, enquanto nas escinas essa ligação ocorre diretamente na aglicona em C-21, o que pode estar relacionado ao fato do **LBL-D** apresentar o grupamento acila e não ser hemolítico.

A análise de componentes principais confirma os resultados obtidos a partir da análise de agrupamentos. Existem analogias que podem ser feitas referentes a parâmetros específicos das referidas técnicas, mas que, muitas vezes, se intercorrelacionam. Apesar das características compartilhadas entre as análises de agrupamentos e de componentes principais, existem diferenças entre as técnicas estatísticas que devem ser mencionadas. A análise de componentes principais não é capaz de discriminar estruturas como a **DHSYA** que possui características estruturais únicas se comparada as demais saponinas derivadas de *G. max*, assim como as variáveis **C3-d** e **C22-carb** que são exclusivas da **DHSYA**. Além disso, a sobreposição dos vetores dificulta a observação de modificações leves que podem refletir aspectos estruturais importantes. Já a análise de agrupamentos é capaz de discriminar estruturas de características únicas e variáveis exclusivas, além de possuir um aspecto gráfico mais estruturado, representado pelos dendrogramas, que permite a visualização de modificações leves. A vantagem da análise de componentes principais em relação à de agrupamentos se refere à possibilidade de analisar o sentido dos vetores, que quando em sentidos opostas, refletem a relação de oposição entre as variáveis, características que não é observada pela análise de agrupamentos.

4.1.4.2 Análise de fatores comuns

No intuito de diminuir a sobreposição dos vetores, observada na análise de componentes principais das observações, principalmente, após a introdução das atividades biológicas as variáveis de entrada, realizou-se a análise de fatores comuns.

A *análise de fatores comuns* tem os mesmos objetivos da *análise de componentes*, possuindo praticamente as mesmas ferramentas com pequenas modificações de nomenclatura, sendo que os *componentes* passam a ser denominados *fatores*, embora o significado do termo se mantenha inalterado. Entretanto, na análise de fatores existe a possibilidade de aplicarmos a *rotação fatorial* que leva a soluções fatoriais geralmente mais claras e simples. Ao se rotacionar a matriz fatorial, redistribui-se a variância dos primeiros fatores para os últimos com o objetivo de atingir um padrão fatorial mais simples e teoricamente mais significativo (HAIR e col., 2005).

Realizaram-se análises das observações após a inclusão das atividades biológicas codificadas **Adj** e **Hemo**, assim como, após a introdução das variáveis **PHA** e **CH50%** as variáveis de entrada para tentar contorna o problema da sobreposição dos vetores. Os resultados referentes ao número de fatores a serem extraídos, podem ser visualizados na figura 9.

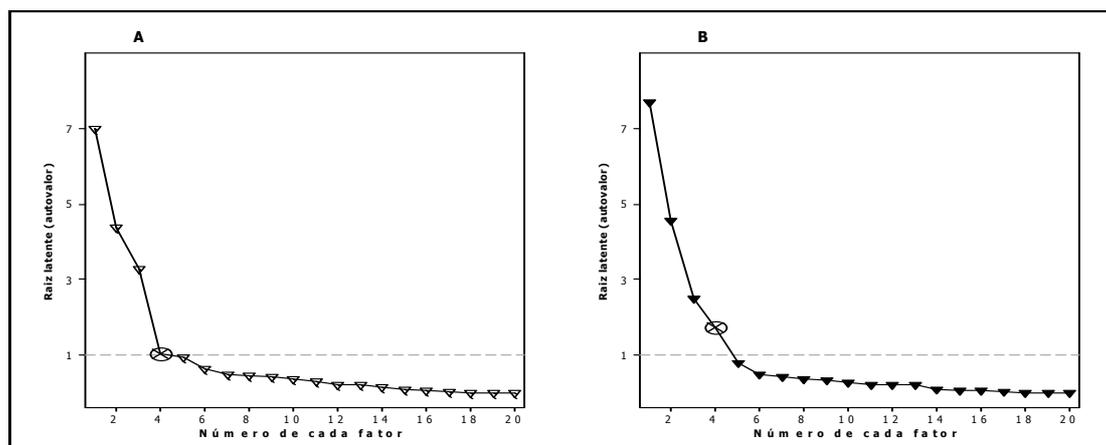


Figura 9. Gráficos de autovalor das observações na (A) presença de atividades biológicas codificadas e na (B) presença de atividades biológicas representadas por variáveis numéricas.

Utilizando o critério da raiz latente, devem ser extraídos quatro fatores nos dois casos, número que é capaz de explicar a maior proporção da variabilidade que engloba as variáveis em análise. Utilizou-se este número de fatores, posteriormente, nas análises das observações na presença de atividades biológicas representadas por

variáveis codificadas e variáveis numéricas. A representação gráfica das análises pode ser visualizada na figura 10.

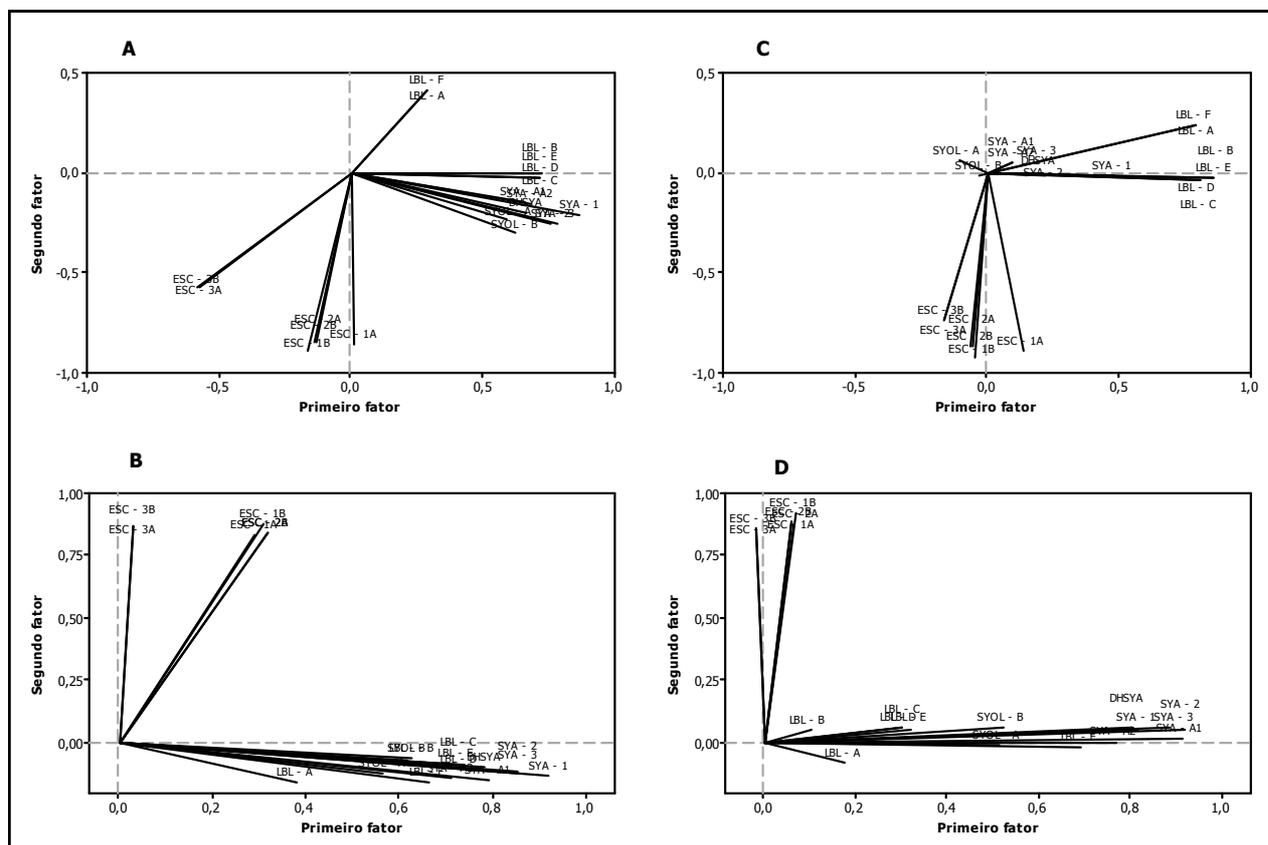


Figura 10. Representação gráfica da análise de fatores comuns das observações na (A e C) presença de atividades biológicas codificadas e na (B e D) presença de atividades biológicas representadas por variáveis numéricas, sendo que, A e B representam a matriz não-rotacionada e C e D representam a matriz rotacionada.

Analisando-se a representação gráfica após a introdução de variáveis biológicas codificadas, no caso da *matriz não-rotacionada* (figura 10A) a estrutura do gráfico não se modifica, ao compararmos com os resultados obtidos pela análise de componentes principais (figura 8A). No entanto, após a *rotação ortogonal de fatores* (figura 10C), observa-se que as saponinas derivadas de *G. max* ocupam uma posição central, já os lablabosídeos se mantêm praticamente na mesma posição que ocupavam anteriormente, assim como as escinas. Tal comportamento indica uma segregação das

saponinas derivadas de soja em relação às derivadas de *D. lablab*, revelando a ocorrência de três grupos distintos na população analisada, como observado na análise de agrupamentos, mas dificultado pela sobreposição dos vetores na análise de componentes principais.

A análise de fatores comuns das observações após a introdução de atividades biológicas representadas por variáveis numéricas (**PHA** e **CH50%**), da *matriz não-rotacionada* (figura 10 B) apresenta a mesma estrutura geral da análise de componentes principais (figura 8C), porém ocorre uma inversão nos quadrantes. No entanto, analisando-se a representação gráfica da *matriz rotacionada* (figura 10D) observa-se que o módulo do vetor, de uma maneira geral, está relacionado ao potencial imunoadjuvante no caso dos lablabosídeos e saponinas de soja. Esta analogia é restrita as saponinas de *G. max* e *D. lablab*, ou seja, na família dos lablabosídeos, o **LBL-F** apresenta o maior potencial imunoadjuvante e, portanto, o maior módulo vetorial. O mesmo ocorre entre as saponinas de soja, sendo que a **SYA-A1** que apresenta a maior potencial imunoadjuvante, também apresenta o maior módulo vetorial. A análise das escinas fica prejudicada em virtude da sobreposição de vetores.

Pela análise de fatores comuns foi possível relacionar novos parâmetros como o módulo dos vetores, que apresentam uma relação importante com a capacidade imunoadjuvante das saponinas de *G. max* e *D. lablab*. Entretanto, a sobreposição dos vetores continua sendo o grande limitador da saída gráfica da análise de fatores comuns, assim como ocorre na análise de componentes principais dificultando a avaliação visual das correlações entre as saponinas. A rotação de certa forma permite uma maior diferenciação entre as saponinas em virtude da diluição da variância comum entre os fatores principalmente os últimos, o que leva a soluções fatoriais geralmente mais simples se comparada à análise de componentes principais.

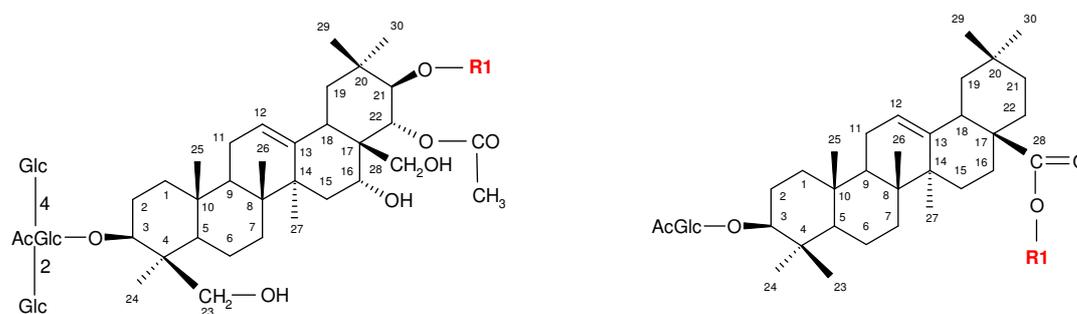
4.1.5 Análise da capacidade preditiva do modelo proposto

As quatro saponinas selecionadas para este teste (tabela 4) fazem parte do trabalho de ODA e col. (2000) e, portanto, os dados referentes à atividade hemolítica e

imunoadjuvante foram obtidos pela mesma técnica. As considerações apresentadas na análise realizada externamente foram coerentes com os parâmetros preditivos de discriminação ou agrupamento, deduzidos do presente trabalho.

Tanto a análise de agrupamentos, como a análise de componentes principais na presença de atividades biológicas representadas por variáveis numéricas (**PHA** e **CH50%**), revelaram que a capacidade hemolítica das saponinas avaliadas pode estar relacionada a presença dos substituintes, **C16-OH**, **C17-metox**, **C22-carbmet** e **C28-OH**, além dos grupamentos acila **C21-tig** e **C21-ang**.

Tabela 4. Representação estrutural das saponinas triterpênicas utilizadas na avaliação da capacidade preditiva do modelo proposto.



Saponina (Abreviação)	R1	PHA*	CH50%**	Saponina (Abreviação)	R1	PHA*	CH50%**
Escina IV (SAP-1)	acetil	61	240	Chikusetsusaponina IVa (SAP-4)	Glc	35	≥600
Escina V (SAP-2)	isobutiril	16	57				
Escina VI (SAP-3)	2 metilbutanoil	16	55				

* Título de PHA obtido mediante teste de Hemaglutinação Passiva;

** Concentração hemolítica 50% (CH50%) em µg/ml;

Glc (Glicose)

Nesse sentido, observa-se que **SAP-1**, **SAP-2** e **SAP-3**, isoladas de *A. hippocastanum*, possuem os substituintes característicos de saponinas hemolíticas, o que foi confirmado pelo valor de **CH50%**, que é menor que 600 µg/ml. Ao avaliar-se a capacidade hemolítica de **SAP-1**, **SAP-2** e **SAP-3**, observou-se que as duas últimas apresentam uma maior capacidade hemolítica e uma menor capacidade imunoadjuvante do que **SAP-1**. A maior capacidade hemolítica pode estar relacionada à

natureza do grupamento acila em C-21, já que as saponinas **SAP-2** e **SAP-3**, apresentam grupamentos acila de maior volume e mais lipofílicos que **SAP-1**. Confirmando os resultados obtidos pela análise componentes principais, pode-se observar que entre **SAP-1**, **SAP-2** e **SAP-3** existe uma relação de oposição entre **PHA** e **CH50%**, ou seja, quanto maior sua capacidade hemolítica, menor sua capacidade imunoadjuvante.

Avaliando-se a **SAP-4**, em virtude de suas características estruturais como, por exemplo, a presença de dois açúcares ligados a aglicona, a presença de **C17-carbose**, e a ausência dos substituintes que determinam a atividade hemolítica (**C16-OH**, **C17-metox**, **C22-carbmet** e **C28-OH**) associada à relação inversa entre as atividades imunoadjuvante e hemolítica, concluiu-se que **SAP-4** tende a apresenta atividade imunoadjuvante e ser não hemolítica. Esta previsão foi confirmada posteriormente pelos seus valores de **PHA** e sua **CH50%**, já que **CH50%** é maior ou igual a 600 µg/ml, indicando que a saponina não apresenta atividade hemolítica, e seu valor de **PHA** é significativamente maior que o do branco (maior que 15), indicando que a saponina apresenta atividade imunoadjuvante.

Os resultados indicam que o modelo proposto apresenta um bom caráter preditivo, ou seja, é capaz de predizer a capacidade imunoadjuvante e hemolítica de uma saponina triterpênica a partir de determinados grupamentos presentes em sua estrutura, que foram, anteriormente, identificados a partir da análise de agrupamentos e de componentes principais. Contudo, a capacidade preditiva do modelo se restringe ao conjunto de saponinas analisadas, que são derivadas principalmente do ácido oleanólico, não podendo ser estendido a toda população de saponinas.

5 CONCLUSÕES

5 CONCLUSÕES

- As técnicas estatísticas multivariadas mostraram-se sensíveis seletivas e específicas na análise estrutural das saponinas triterpênicas, fornecendo meios gráficos para o estabelecimento de relações estruturais envolvidos com as atividades imunoadjuvante e hemolítica das saponinas selecionadas.
- A análise de agrupamentos mostrou-se mais sensível que as análises de componentes principais e fatores comuns, ao permitir a visualização de pequenas diferenças nos dendrogramas que refletem importantes aspectos estruturais, além de discriminar estruturas e variáveis com características peculiares, como é o caso de Deidrosojasaponina (**DHSYA**) e suas variáveis exclusivas(**C3-d** e **C22-carb**).
- A análise de componentes principais permitiu a visualização de relações de oposição entre as variáveis, demonstradas por meio dos sentidos opostos dos vetores, mas que não são observadas na análise de agrupamentos.
- A análise de fatores comuns propiciou soluções fatoriais mais simples e significativas que as obtidas mediante análise de componentes principais devido à rotação ortogonal.
- Na análise de componentes principais e, em menor grau, na análise de fatores comuns, a sobreposição dos vetores limitou a visualização de pequenas modificações que refletem importantes aspectos estruturais.
- A utilização das diferentes técnicas de análise multivariada, em associação propiciou uma extração mais efetiva dos parâmetros estruturais relacionados às atividades imunoadjuvante e hemolítica.

Conclusões

- Os diferentes métodos aplicados revelaram que a atividade imunoadjuvante pode estar relacionada à presença de uma hidroxila em C-21 de uma metila em C-17, e de dois açúcares ou cadeia de açúcares ligados a aglicona. Já a atividade hemolítica parece estar relacionada à presença de uma hidroxila em C-16, de um CH₂OH em C-17, de uma acetila em C-22 e de um grupamento acila em C-21.
- A análise de componentes principais revelou a existência de uma relação de oposição entre **PHA** e **CH50%**, assim como entre **Adj** e **Hemo**, indicando que as atividades imunoadjuvante e hemolítica apresentam uma relação inversa.
- O modelo preditivo proposto foi capaz de prever o potencial imunoadjuvante e hemolítico das saponinas, mediante avaliação da relação entre os diferentes substituintes ligados a aglicona e as atividades imunoadjuvante e hemolítica.

6 PERSPECTIVAS

6. PERSPECTIVAS

- O presente trabalho demonstrou que as técnicas multivariadas podem ser utilizadas na análise estrutural de saponinas como também para outras substâncias. Nesse sentido, pretende-se expandir essa metodologia as saponinas que possuem outras atividades biológicas, como por exemplo, antiviral e antiinflamatória, e assim buscar relações estruturais relacionadas a essas atividades.
- A capacidade preditiva do método pode ser utilizada para guiar a escolha de saponinas a serem avaliadas quanto ao potencial imunoadjuvante em função da presença de determinados substituintes, identificados mediante técnicas multivariadas, que podem ser importantes para a atividade imunoadjuvante.
- Pretende-se avaliar os substituintes relacionados com as atividades imunoadjuvante e hemolítica, quanto a sua polaridade, a sua disponibilidade eletrônica e outros parâmetros que reflitam características físico-químicas dos grupamentos, para se determinar a influência destes sobre as referidas atividades biológicas.

7 REFERÊNCIAS

7 REFERÊNCIAS

ASCARATEIL S.; DUPUIS L. Surfactants in vaccine adjuvants: description and perspectives. **Vaccine**, v. 24S2, p. S2/83 – S2/85, 2006.

AUCOUTURIER J.; DUPUIS L.; GANNE V. Adjuvants designed for veterinary and human vaccines. **Vaccine**, v. 19, p. 2666–2672, 2001.

ESTRADA A.; KATSELIS G. S.; LAARVELD B.; BARL B. Isolation and evaluation of immunological adjuvant activities of saponins from *Polygala senega* L. **Comparative Immunology, Microbiology & Infectious Diseases**, v. 23, p. 27 – 43, 2000.

FLECK J. D.; KAUFFMANN C.; SPILKI F.; LENCINA C. L.; ROEHE P. M.; GOSMANN G. Adjuvant activity of *Quillaja brasiliensis* saponins on the immune responses to bovine herpesvirus type 1 in mice. **Vaccine**, v. 24, p. 7129 – 7134, 2006.

GLAUERT, A.M.; DINGLE, J.T.; LUCY, J.A. Action of saponin on biological cell membranes. **Nature**, v.196, p.952-955, 1962.

HAIR J. F.; ANDERSON R. E.; TATHAM R. L.; BLACK W. C. **Análise multivariada de dados**. 5 ed. Porto Alegre, Bookman, 2006.

HOSTETTMANN, K.; MARSTON,A. **Saponins – chemistry and pharmacology of natural products**. Cambridge: University, 1995.

HUNTER R. Overview of vaccine adjuvant: present and future. **Vaccine**, v. 20S3, p. S7–S12, 2002.

KARABALIEV, M.; KOHEV, V. Interaction of solid supported thin lipid films with saponin. **Sensors and Actuators B: Chemical**, v.88, p.101-105, 2003.

MATSUDA, H.; LI Y.; MURAKAMI T.; NINOMIYA, K; YAMAHARA, J.; YOSHIKAWA M. Effects of escins Ia, Ib, IIa and IIb from horse chestnut, the seeds of *Aesculus hippocastanum* L., on acute inflammation in animals. **Biological & Pharmaceutical Bulletin**, v.20, p. 1092 – 1095, 1997.

ODA K.; MATSUDA H.; MURAKAMI T.; KATAYAMA S.; OHGITANI T.; YOSHIKAWA M. Adjuvant and Haemolytic Activities of 47 Saponins Derived from Medicinal and Food Plants. **Biological Chemistry**, v. 381, p. 67 – 74, 2000.

ODA K.; MATSUDA H; MURAKAMI T; KATAYAMA S.; OHGITANI T.; YOSHIKAWA M. Relationship between adjuvant activity and amphipathic structure of soyasaponins. **Vaccine**, v.21, p. 2145 – 2151, 2003.

O' HAGAN, D.T.; MACKINCHAN, M.L.; SINGH, M. Recent developments in adjuvants for vaccines against infectious diseases. **Biomolecular Engineering**, v 18, p. 69–85, 2001.

OLESZEK W.; MARSTON A. **Saponins in food, feedstuffs and medicinal plants**. 1. ed. Boston: Kluwer Academic Publishers Dordrecht, London: Health & Fitness, 2000.

SCHENKEL, E. P.; GOSMANN, G.; ATHAYDE, M. L. Saponins In: SIMÕES, C. M. O.; SCHENKEL, E. P.; GOSMANN, G; MELLO, J. C. P.; MENTZ, L. A.; PETROVICK, P. R. (org) **Farmacognosia: da planta ao medicamento**. 5 ed. rev. ampl. Porto Alegre: UFRGS, Florianópolis: UFSC, 2003. Cap. 27, p. 711 – 740.

STEURER S.; WURGLICS M.; LIKUSSAR W.; BURMISTROV K.; MICHELITSCH A.; SCHUBERT-ZSILAVECZ M. Lack of correlation between surface and interfacial activities of saponins and their hemolytic properties. **Pharmazie**, v. 54, p. 766 – 767, 1999.

Referências

SÜLEYMAN H.; MSHVILDADZE V.; GEPIREMEN A.; ELIAS R. Acute and chronic antiinflammatory profile of the ivy plant, *Hedera helix*, in rats. **Phytomedicine**, v.10, p. 370 – 374, 2003.

SUN H. X.; QIN F.; YE Y.P. Relationship between haemolytic and adjuvant activity and structure of protopanaxadiol-type saponins from the roots of *Panax notoginseng*. **Vaccine**, v. 23, p. 5533 – 5542, 2005.

SUN H. X.; PAN H. J. Immunological adjuvant effect of *Glycyrrhiza uralensis* saponins on the immune responses to ovalbumin in mice. **Vaccine**, v. 24, p. 1914 – 1920, 2006.

SUN H. X. Adjuvant effect of *Achyranthes bidentata* saponins on specific antibody and cellular response to ovalbumin in mice. **Vaccine**, v. 24, p. 3432 – 3439, 2006.

SUN J.; HU S.; SONG X. Adjuvant effects of protopanaxadiol and protopanaxatriol saponins from ginseng roots on the immune responses to ovalbumin in mice. **Vaccine**, v. 25, p. 1114 – 1120, 2007.

TAKECHI M.; TANAKA Y. Haemolytic time course differences between steroid and triterpenoid saponins. **Planta Medica**, v. 61, p. 76 – 77, 1995.

WANG H. X., NG T. B. Natural products with hypoglycemic, hypotensive, hypocholesterolemic, antiatherosclerotic and antithrombotic activities. **Life Sciences**, v.65, p. 2663 – 2677, 1999.