

RAE-CEA 12P14

“Estudo do papel da proteína Galectina-3 no fluxo metabólico de melanomas cutâneos murinos”.

Profa. Júlia Maria Pavan Soler

Bruno César Galhardo

Daniel Walter Martins

São Paulo, Julho de 2012

**RELATÓRIO DE ANÁLISE ESTATÍSTICA –
CENTRO DE ESTATÍSTICA APLICADA**

TÍTULO:

*“Estudo do papel da proteína Galectina-3 no fluxo metabólico de
Melanomas cutâneos murinos”*

PESQUISADOR: *Antônio Carlos Bloise Jr.*

ORIENTADOR: *Prof. Said Rahnamaye Rabbani*

INSTITUIÇÃO: *Instituto de Física da Universidade de São Paulo - IFUSP*

ORIGEM DO ESTUDO: *Ciências Biológicas*

RESPONSÁVEIS PELA ANÁLISE:

Profa. Júlia Maria Pavan Soler

Bruno César Galhardo

Daniel Walter Martins

REFÊRENCIA DESTE TRABALHO:

SOLER, J.P., RABBANI, S.R., BLOISE, A.C.

Relatório de Análise Estatística sobre o projeto : **“Estudo do papel da Galectina-3 no fluxo metabólico de Melanomas cutâneos murinos”**. São Paulo, IME-USP, 2012. (RAE-CEA – 12P14)

FICHA TÉCNICA

REFERÊNCIAS BIBLIOGRÁFICAS:

JOHNSON, R. A. and WICHERN, D. W., Applied Multivariate Statistical Analysis, Sixth Edition, Pearson Education, Inc., 2007.

BUSSAB, W.O. e MORETTIN, P. A., Estatística Básica. São Paulo: Saraiva, 2002. 5^a edição

KUTNER, M.H., NACHTSHEIM, C.J.; NETER, J. e LI, W. (2005), Applied Linear Statistical Models, Irwin

AITCHISON, J. (1986,) The Statistical Analysis of Compositional Data, Chapman & Hall, The Blackburn Press.

HAIR, J. F., BABIN, W. C. (2006), Multivariate data analysis. 6th ed. New Jersey: Prentice Hall.

PROGRAMAS COMPUTACIONAIS UTILIZADOS:

SAS 9.3 for Windows®, SAS Institute Inc, versão 6.1.7601

Microsoft Excel for Windows®, versão 2010.

Microsoft Word for Windows®, versão 2010.

R for Windows®, versão 2.15.2.

Minitab for Windows, Minitab Inc, versão 16

TÉCNICAS ESTATÍSTICAS UTILIZADAS: (entre parênteses a classificação ISI)

Análise Descritiva – Medidas de Posição, Associação e Gráficos (03:020)

Análise de Componentes Principais (06:070)

Análise Descriminatória (06:050)

Análise de Variância (08:010)

Regressão Logística (07:090)

ÍNDICE

1. Introdução.....	6
2. Objetivos	6
3. Descrição do Estudo	8
4. Análise Descritiva dos Dados.....	10
4.1 Análise Descritiva Não Supervisionada	10
4.2 Análise Descritiva Supervisionada.....	13
4.3 Transformação de Aitchison	13
5. Análise Inferencial	15
5.1 Análise de Componentes Principais	15
5.2 Análise de Variância	16
5.3 Análise Discriminante com Grupos de Metabólitos.....	20
5.4 Regressão Logística	21
6. Apêndice	26
Apêndice 1. Tabelas e Gráficos Descritivos	26
Apêndice 2. Componentes Principais	33
Apêndice 3. Análise de Variância	39
Apêndice 4. Análise Discriminante	79
Apêndice 5. Regressão Logística	81

RESUMO

O estudo realizado trata de melanoma, um tipo específico de câncer de pele. As unidades amostrais são 32 células de melanoma, os melanócitos, submetidos a diferentes condições de cultivo: Linhagem (G2:Galectina e N3), Oxigênio (Hipóxia e Normóxia) e Glicose (Controle e Alta). As variáveis respostas são a quantidade de cada um de 31 metabólitos (resíduos da respiração celular) avaliados em cada célula, expressa em forma de porcentagem. O objetivo do estudo é compreender o impacto da expressão da proteína Galectina-3 no metabolismo de células tumorais e encontrar padrões metabólicos gerais, associando a etapas do desenvolvimento tumoral.

As abordagens estatísticas adotadas foram construção de gráficos *Box Plot*, *Dot plot* e tabelas descritivas para observar o comportamento dos metabólitos; em seguida, após utilizar a transformação de Aitchison para Dados Composicionais, partiu-se para análises marginais, de cada variável separadamente, via ANOVA e seleção dos metabólitos mais importantes na segmentação dos fatores. Por fim, as Análises Discriminante e Regressão Logística, com as variáveis selecionadas pela ANOVA, construíram critérios para classificar as amostras segundo às condições experimentais.

As conclusões obtidas apontam que é possível criar diferentes grupos de metabólitos: os que não diferenciam nenhum dos três fatores de cultivo estudados, alguns outros que segmentam somente os efeitos principais destes fatores, e os demais grupos que contém interações. A partir desses grupos, que correspondem a seleção de metabólitos (variáveis), as abordagens multivariadas de segmentação puderam mostrar suas efetividades e apresentar resultados de diferenciação, por exemplo, a taxa de acerto da discriminação entre os níveis do fator Linhagem, obtida através da análise de Regressão Logística, é 87,5%.

1. Introdução

Câncer é o nome dado a um conjunto de mais de 100 doenças, as quais têm em comum o crescimento desordenado de células que invadem os tecidos e órgãos, podendo espalhar-se rapidamente para outras regiões do corpo. Dividindo-se de forma acelerada, estas células tendem a ser muito agressivas e incontroláveis, determinando a formação de tumores, isto é, acúmulo de células cancerígenas. O câncer de pele, em particular, origina-se por intermédio do desenvolvimento desordenado de células da pele, que pode ter origem por diversas causas como, por exemplo, grandes períodos de exposição aos intensos raios ultravioleta.

O melanoma é um tipo de câncer de pele, conhecido por ser o mais grave. Sua ocorrência é mais rara, porém com alta letalidade e tem motivado muitas pesquisas a seu respeito. O evento motivador para este estudo foi a observação de que em algumas células de melanoma cutâneo em camundongos houve a inibição de uma proteína chamada Galectina-3, levando a célula à morte, pois a Galectina-3 atua no controle do processo de respiração e metabolismo celular.

Tem-se conhecimento de que alguns compostos químicos advindos do processo de respiração, os Metabólitos, permanecem no interior da célula após o metabolismo e, no contexto de células infectadas pelo melanoma, deseja-se entender se há relação entre a quantidade de cada um desses resíduos respiratórios com as características de inibição ou não da proteína Galectina-3.

2. Objetivos

O objetivo deste projeto é identificar as variações nas quantidades de metabólitos presentes em células de melanoma cutâneo murino (camundongos) e sua relação com a galectina-3, permitindo que uma nova célula possa ser classificada com relação aos três fatores de controle do estudo (Linhagem, Glicose e Oxigênio), de acordo com os valores observados dos metabólitos importantes na classificação. Desta maneira, a nova célula será associada a uma das combinações dos níveis dos fatores do estudo, sendo assim, a classificação desta unidade amostral trará informações importantes sobre o gene Galectina-3, permitindo extrações sobre seu silenciamento.

3. Descrição do Estudo

A obtenção dos dados passa inicialmente pelo processo de seleção de células de camundongo. As células disponíveis para o estudo encontravam-se congeladas e foram retiradas de um mesmo animal e mantidas em baixa temperatura para fins de conservação. De forma mais específica, foram extraídas células saudáveis do tecido epitelial dessa cobaia e se iniciou o processo de indução ao tumor, isto é, uma exposição à outra célula já contaminada pelo Melanoma e, então, as novas passam também a desenvolver a doença.

Nesse grupo de células verificou-se que a proteína Galectina-3 teve sua ação inibida devido à manifestação da moléstia. Então, a partir desse grupo foram replicadas 32 outras células, em que algumas passaram por um processo de alteração no DNA e voltaram a produzir a Galectina-3, a saber, Linhagem G2, e as restantes continuaram sem expressão da proteína, Linhagem N3.

Células de cada Linhagem (G2 e N3) foram expostas a meios de cultura com variações do Nível de Oxigênio (Normóxia e Hipóxia) e ao Nível de Glicose (Controle e Alta). Deste modo, para o estudo tem-se 7 grupos definidos de acordo com combinações dos três fatores, Linhagem, Oxigênio e Glicose (Tabela 1). Note que a condição N3, Hipóxia e Glicose Alta não entrou para o estudo.

Tabela 1 : Grupos de estudo com os correspondentes tamanhos amostrais

LINHAGEM	OXIGÊNIO	GLICOSE	TAMANHO AMOSTRAL
G2	NORMÓXIA	CONTROLE	6
G2	NORMÓXIA	ALTA	7
G2	HIPÓXIA	CONTROLE	3
G2	HIPÓXIA	ALTA	6
N3	NORMÓXIA	CONTROLE	4
N3	NORMÓXIA	ALTA	3
N3	HIPÓXIA	CONTROLE	3
N3	HIPÓXIA	ALTA	0

Há, portanto, 7 combinações resultantes dos 3 fatores, totalizando um tamanho amostral igual a 32. Estas células passaram por um processo de filtragem para eliminação da influência de outros fatores que não sejam o do objetivo do estudo e, em seguida, utilizou-se um método de separação de resíduos por ressonância magnética para identificar o conjunto de metabólitos. Com o auxílio de um espectrograma, foi avaliada a variável resposta para o estudo que é a **quantidade (porcentagem) de cada metabólito presente nas células submetidas a diferentes condições experimentais**.

A obtenção dessas quantidades advém de um processo sofisticado, pois visto que se trata de célula e resíduos menores que ela, cujos tamanhos estão na ordem dos nanômetros, instrumentos e métodos usuais não dariam conta de coletar estas medidas. Então, para fazer a mensuração e identificação de quais são os compostos (metabólitos) presentes na célula, avalia-se no espectrômetro (ver Figura 1) alguns picos de frequência característicos e, com base em conhecimentos específicos, o pesquisador identifica qual é o metabólito associado ao pico. Depois de identificado, para quantificar o metabólito, calcula-se a área sob a curva padronizada do espectrograma e esta área sob a curva representa a quantidade percentual daquele metabólito em relação a todos os metabólitos presentes na célula. Tem-se então que os conceitos de área sob a curva e quantidade do metabólito serão tratadas como iguais na nomenclatura deste relatório.

Dentre todos os metabólitos existentes na célula, considerar-se-ão apenas 31 para este estudo, cujos nomes estão especificados com os seguintes códigos P1, P2,..., P31. O 32º, Poutros, que é a presença de outros metabólitos senão os primeiros, fechando então a soma das porcentagens em 100% para cada célula avaliada, que é a unidade amostral do estudo.

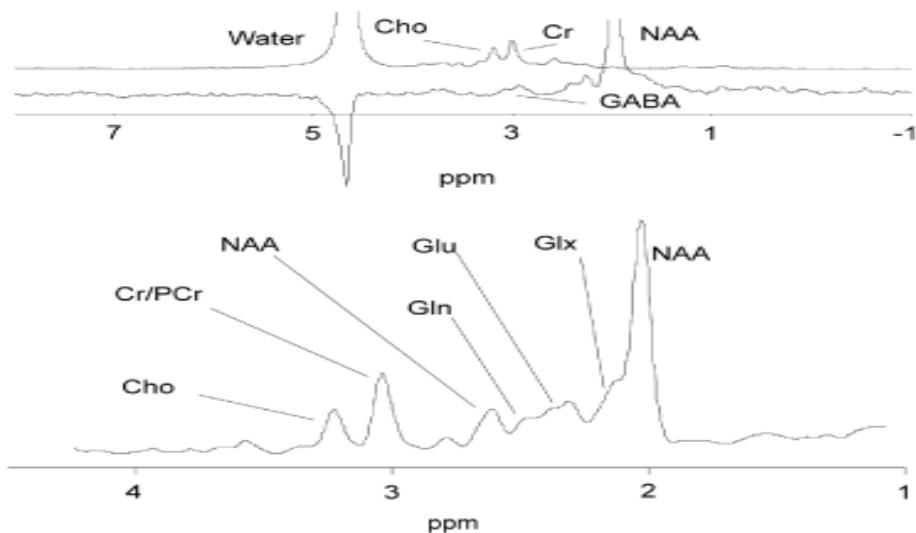


Figura 1 : Ilustração de um Espectrograma

4. Análise Descritiva dos Dados

Os dados sob estudo podem ser dispostos em uma matriz com 32 linhas (as unidades amostrais, células) e 32 colunas (definindo as variáveis, metabólitos), tendo a quantidade de cada metabólito (em %) como variável resposta. As unidades amostrais estão agrupadas em 7 grupos definidos pela combinação dos três fatores sob estudo: Linhagem, Oxigênio e Glicose. Vale ressaltar que para a realização de análises multivariadas, considerando os 32 metabólitos conjuntamente, o tamanho amostral (32 unidades amostrais) pode não ser adequado. Uma recomendação da literatura é que o número de amostras deve ser da ordem de 5 vezes o número de variáveis (HAIR, 2006).

Além disso, não há cruzamento de todos os níveis dos fatores Linhagem, Oxigênio e Glicose, não sendo possível analisar eventuais interações entre esses fatores. Outro ponto de atenção se deve ao fato de que o experimento não é balanceado em todos os cruzamentos dos fatores de controle, ou seja, não há o

mesmo número de réplicas para todas as combinações.

Na análise descritiva dos dados foram construídos gráficos Box Plot, Dot Plot e tabelas com estatísticas descritivas para observar o comportamento dos metabólitos. Maiores detalhes sobre a construção de tais gráficos e cálculos envolvidos podem ser encontrados em BUSSAB e MORRETIN (2002).

4.1 Análise Descritiva Não Supervisionada

Na Análise Descritiva que denominamos não supervisionada foram desconsideradas as combinações dos fatores sob estudo, ou seja, essa análise visa encontrar possíveis semelhanças e características entre as unidades amostrais, sem considerar os grupos a qual elas pertencem. Assim, a formação dos grupos sob as condições iniciais de cultivo experimental pode ser corroborada, ou contrariada, dependendo do comportamento dos resultados encontrados nessa análise não supervisionada.

4.1.1 Tabela Descritiva

Inicialmente construiu-se a Tabela de Medidas Descritivas (Tabela 1.1, do Apêndice 1), com as seguintes medidas resumo: Média, Desvio Padrão, Mínimo, Máximo e o Coeficiente de Variação. Analisando as médias por metabólito, nota-se que os dados têm uma ordem de grandeza pequena (10^{-2}) e existem muitos metabólitos semelhantes, em relação ao baixo valor médio. A variável Poutros, que visa somente à complementaridade das proporções, reteve muito das quantidades totais, entre 35,40% e 53,53%. A respeito da dispersão das respostas, preferiu-se considerar o Coeficiente de Variação (CV) como medida, pois é uma medida adimensional e mostra a relação direta entre a variabilidade dos dados em relação à média.

Destaca-se, o metabólito P28, que teve CV igual a 195,85%, causado por

três observações muito atípicas, sendo que uma delas chegou a ocupar 25,94% dos metabólitos encontrados na célula. Para os outros 31 metabólitos, observaram-se coeficientes de variação oscilando entre 8% e 95%, indicando que a dispersão dos dados, relativamente à média, é de baixa a moderada.

Devido à alta variabilidade e ao desconhecimento da causa que gerou observações atípicas, a análise descritiva do Metabólito P28 foi feita com os valores atípicos e na ausência deles. Ver Tabelas : 1.1 e 1.2 no Apêndice 1.

4.1.2 - Box Plot

Com a finalidade de avaliar as medidas apresentadas na tabela descritiva de uma maneira gráfica, construiu-se o *Box Plot (Gráfico 1)* dos valores das porcentagens para todos os metabólitos, ordenando-os de forma crescente pela média. A variável Poutros não foi incluída por questões de distorção de escala.

Destaca-se, inicialmente, a variabilidade do metabólito P28 (ver gráfico 1), com uma observação muito atípica. Para os outros metabólitos, nota-se que a dispersão aumenta proporcionalmente ao valor médio, e, em alguns casos, apresentam *outliers*, observações que se distanciam das demais. O gráfico que não contém a observação atípica do Metabólito P28 encontra-se no Apêndice 1(gráfico 1.5).

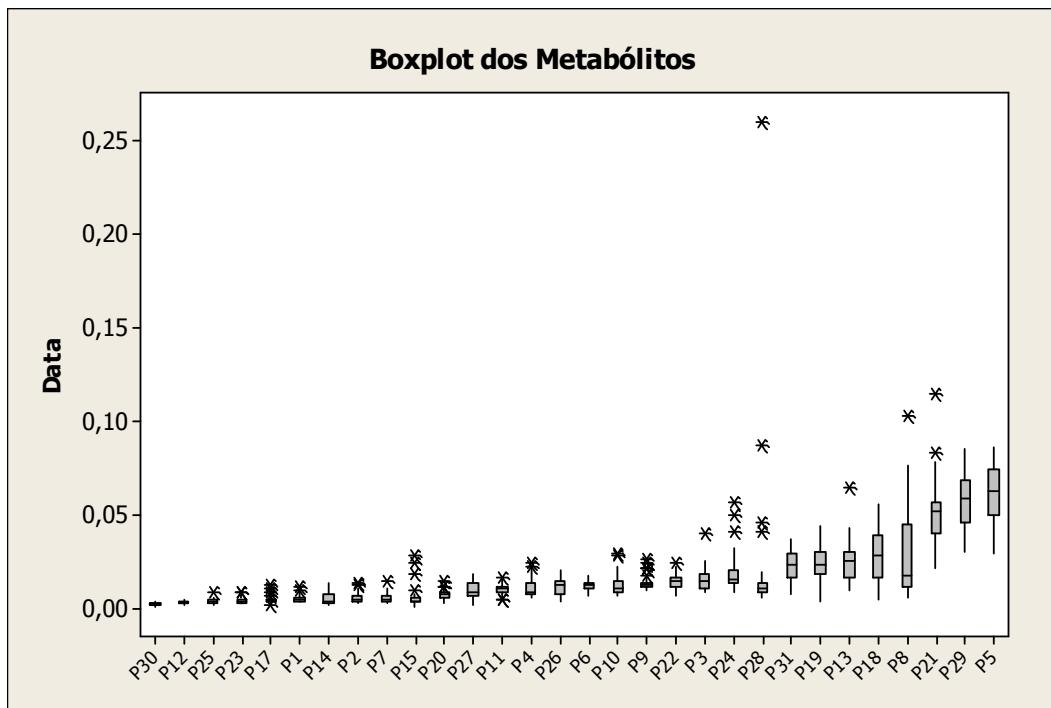


Gráfico 1: Box Plot dos 31 Metabólitos (sem Poutros)

4.2 Análise Supervisionada

Diferentemente da abordagem anterior, na Análise que denominamos de Supervisionada consideram-se as condições experimentais das unidades amostrais, então, as observações são estudadas de acordo com as combinações de Linhagem, Oxigênio e Glicose. O objetivo desta análise é encontrar padrões de metabólitos que diferenciem e caracterizem as células, segundo as 7 combinações disponíveis:

4.2.1 DOT PLOT

O Dot Plot é um recurso gráfico poderoso em situações que se têm poucos dados e se deseja visualizar o comportamento dos mesmos, seja

comparativamente a grupos ou individualmente, olhando um único grupo. Na análise supervisionada dos dados, as 7 combinações disponíveis dos fatores Linhagem, Oxigênio e Glicose foram colocadas em um único gráfico e procurou-se encontrar algum padrão que as caracterizassem. No entanto, a distribuição das observações no gráfico não sugerem um padrão específico com relação às combinações Linhagem x Oxigênio x Glicose.

Na tentativa de facilitar a visualização de algum padrão, considerar-se-ão, então, as amostras divididas em níveis principais dos fatores Linhagem, Oxigênio e Glicose. Ver no Apêndice 1 (Gráficos 1.6, 1.7 e 1.8, sem a observação atípica de P28, Gráficos 1.2, 1.3 e 1.4, com a observação atípica). Contudo, ainda assim, não foi possível encontrar metabólitos que distinguissem visualmente os níveis de linhagem (G2 e N3), Oxigênio (Normoxia e Hipoxia) e Glicose (Controle e Alta).

4.3 Transformação de Aitchison

Dados Composicionais são aqueles que apresentam informações relativas, isto é, partes de um todo. O principal atributo deste tipo de dados é que sua soma é sempre constante e fixa, por exemplo, igual a 1 quando proporções ou igual a 100 quando porcentagens. Visto que os dados deste estudo correspondem proporções relativas a um todo, é apropriado considerar esta estrutura Composicional na análise dos mesmos. Este tipo de dados tem características específicas que devem ser respeitadas na definição de medidas de distância entre observações ou correlações entre as variáveis.

A estrutura matricial é útil para representar tais dados, com as linhas da Matriz Composicional sendo representada pelas unidades amostrais, ao passo que as colunas representam as partes, categorias ou variáveis. Uma das propriedades destes dados é a soma dos valores das variáveis ser constante para cada unidade amostral. As aplicações das análises convencionais, em geral, não são adequadas para examinar o comportamento dos Dados Composicionais. A matriz de correlação entre os metabólitos é apresentada na Tabela 1.2 do Apêndice 1.

Um problema imediato de não tratar esses dados como compostacionais seria detectado por intermédio da análise de correlação entre os metabólitos, onde pelo menos uma das correlações entre as variáveis deve ser negativa. Em termos práticos, quando a proporção de um metabólito aumentar, obrigatoriamente, a proporção de pelo menos um dos demais deve diminuir. Com isso o Coeficiente de Correlação perde sua característica básica: a liberdade de variar entre -1 e +1. Trata-se, portanto, de uma correlação espúria de tendência negativa.

Realizar análise para dados compostacionais por intermédio de métodos estatísticos factícios pode induzir a conclusões equivocadas. Uma alternativa para contornar eventuais problemas e que será abordada neste trabalho é a transformação dos Dados Compostacionais proposta por Aitchison(1983):

Transformação de Aitchison:

$$Y_{ij} = \log [X_{ij} / g(X_j)]$$

Em que:

i : índice de unidade amostral ($i = 1, 2, \dots, n$), $n=32$

j: índice de variável($j=1, 2, \dots, p$), com $p=32$ (31 metabólitos, mais Poutros)

Y_{ii} : Dados Transformados

X_j: Vetor de Observações da j-ésima unidade amostral

X_{ij} : j-ésima observação da i-ésima unidade amostral

g(X_i): Média Geométrica do Vetor de observações **X_i**;

tal que : $g(X_i) = (\prod_{j=1}^p X_{ij})^{1/p}$.

Uma ressalva importante dessa transformação é que ela mantém a ordenação dos dados, mas corrige a distância entre as observações e a correlação entre as variáveis. Tal correção se dá por incorporar a restrição de que os dados são compostacionais. Se uma célula cultivada em Glicose Alta possui mais do Metabólito P5, por exemplo, então nos dados transformados essa ordenação será respeitada e mantida.

5. Análise Inferencial e Análise Multivariada

5.1- Análise de Componentes Principais

Para os dados transformados, partiu-se-a para uma abordagem multidimensional com relação aos metabólitos: Análise de Componentes Principais (ACP). Pela ACP, ao invés de se considerar as respostas para os 32 metabólitos, escolher-se-ão combinações lineares delas, isto é, ponderações entre as 32 respostas que tragam o máximo da informação da variabilidade contida nos dados originais. Em alguns casos, essas combinações podem ser vistas como índices, ou escores, dependendo da natureza das respostas, porém, nem sempre há uma interpretação bem definida das componentes. Neste estudo, as componentes são vistas como contrastes entre as quantidades transformadas desses 32 metabólitos e se utilizaram quatro destes contrastes por motivos de explicação da variabilidade total próxima de 70%.

Há duas abordagens possíveis para a ACP: utilizando matriz de covariâncias ou a matriz de correlações.

Matriz Covariância – Sob este enfoque, as respostas para os 32 metabólitos são consideradas heterocedásticas e esta diferença entre as variâncias é importante na análise. Como consequência, metabólitos com maior variabilidade na resposta transformada tendem a ter pesos maiores nas componentes principais e, de acordo com o gráfico *Box Plot* visto na seção 4.1.2, espera-se que nas componentes principais haja peso alto para os últimos metabólitos, pois eles têm amplitudes de variação maiores. Para optar por esta análise, cabe ao pesquisador averiguar se a natureza dessas variabilidades maiores é importante em relação às menores e qual o significado disto concernente aos metabólitos.

Matriz Correlação - Partindo-se da premissa que os dados são oriundos de escalas diferenciadas, utilizam-se as quantidades de metabólito padronizadas, com o intuito de eliminar a influência de escala, tornando-as variáveis com média 0 e variância 1. Assim, após esta normalização, têm-se homogeneidade entre as respostas e desta maneira os metabólitos contribuem de maneira igualitária nos

pesos das componentes principais.

Utilizando ambas as abordagens e com o objetivo de tentar encontrar uma representação, em duas dimensões, que conseguisse segmentar os 7 grupos formados pelas combinações dos fatores de estudo, buscou-se utilizar o gráfico de dispersão dos escores das componentes principais para facilitar esta visualização. Os resultados obtidos, via matriz de correlação, apontam que os grupos estão todos confundidos entre si, não havendo distinção de nenhum grupo específico com relação aos outros, como se pode observar nos gráficos (2.4,2.5,2.6,2.7,2.8 e 2.9) do Apêndice 2. Sendo assim, uma nova tentativa de segmentação foi adotada pela divisão dos dados somente em efeitos principais dos fatores, isto é, para alguns destes mesmos gráficos procurou-se fazer distinção somente de Linhagens, G2 e N3, e verificar se houve segmentação, depois para Oxigênio, Normóxia e Hipóxia, e, por último, para o fator Glicose. Foi feita uma validação da capacidade de distinção das componentes principais via Análise Discriminante, porém, nenhum desses resultados foi satisfatório para fins de segmentação dos grupos, o que sugere a necessidade de se aplicar outras abordagens estatísticas para se fazer inferências sobre a influência dos metabólitos para classificação dos fatores.

5.2 ANOVA – Análise de Variância

Tendo os dados transformados como entrada, o objetivo da aplicação da ANOVA é encontrar metabólitos que, individualmente, diferenciam as características dos grupos, de forma que fiquem claros quais efeitos dos fatores estão presentes dentro daquele metabólito e, por meio destes efeitos, poder agrupar os metabólitos em classes de efeitos similares, por exemplo, aqueles que só dependem do Oxigênio (com médias diferentes para Normóxia e Hipóxia).

A ANOVA é uma técnica de comparação de médias de distribuições Normais de mesma variância (ver, Neter et al.,2006). O modelo adotado neste estudo na análise dos dados é apresentado no Apêndice 3.

Para os dados transformados, foram feitas então 32 ANOVA's, uma para cada metabólito, seguindo um modelo estrutural como descrito no Apêndice 3, porém, em alguns casos, a suposição de igualdade de variâncias das 7 combinações dos fatores não foi verificada, impossibilitando a abordagem via modelo clássico de ANOVA. Contudo, a utilização da técnica de Mínimos Quadrados Ponderados (MQP), como alternativa à abordagem clássica da ANOVA clássica, tem o poder de corrigir esta heterocedasticidade e ajustar um modelo que envolva os mesmos componentes da ANOVA.

Após as correções necessárias via MQP, todos os modelos ficaram bem ajustados, com as suposições verificadas e validadas. Os gráficos de diagnóstico podem ser encontrados no Apêndice 3(Gráficos de 3.1 a 3.32). A Normalidade pode ser vista pelo gráfico QQ-Plot, indicando que quanto mais próximo de uma reta, melhor a adequação de normalidade, bem como o histograma. A homocedasticidade e independência dos resíduos verificam-se pela análise dos gráficos de dispersão dos resíduos pelos valores ajustados e pela ordem das observações, respectivamente. A partir do ajuste, passa-se a analisar a significância (valor-p < 5%) dos efeitos de interação entre os fatores e, caso estes não sejam significantes, os níveis descritivos (valores-p) dos efeitos principais dos fatores são avaliados. Finalmente, para alguns metabólitos nenhum dos efeitos se mostrou significante, podendo estes serem reduzidos à resposta média, independente dos fatores. As tabelas do Apêndice 3 mostram os resultados agrupando os metabólitos de acordo com os efeitos significantes da ANOVA.

Vejamos dois exemplos em que o modelo mostrou efeito significante apenas para o fator Oxigênio, metabólitos P6 (valor-p=0,001) e P24 (valor-p=0,018). Neste caso, há interesse em entender como as médias destes metabólitos variam para Normoxia e Hipoxia. As extrapolações para os outros fatores são análogas.

O Gráfico 2 apresenta os perfis de médias do metabólito P6 de acordo com Oxigênio. Conclui-se que a média deste metabólito na condição de Hipoxia é significantemente menor que a média em Normoxia (valor-p=0,001).

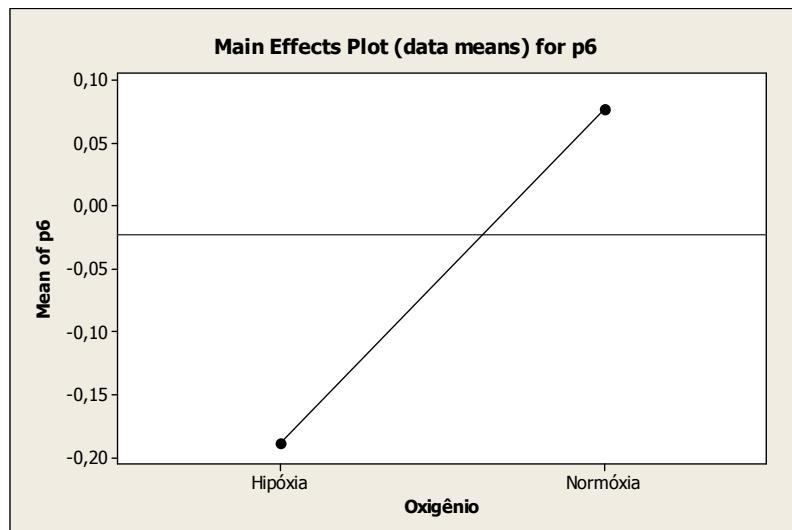


Gráfico 2: Perfis de Médias para a variável transformada de P6.

O Gráfico 3 apresenta os perfis de médias do metabólito P24 de acordo com Oxigênio. Conclui-se que a média deste metabólito na condição de Hipóxia é significantemente maior que a média em Normóxia (valor-p=0,018).

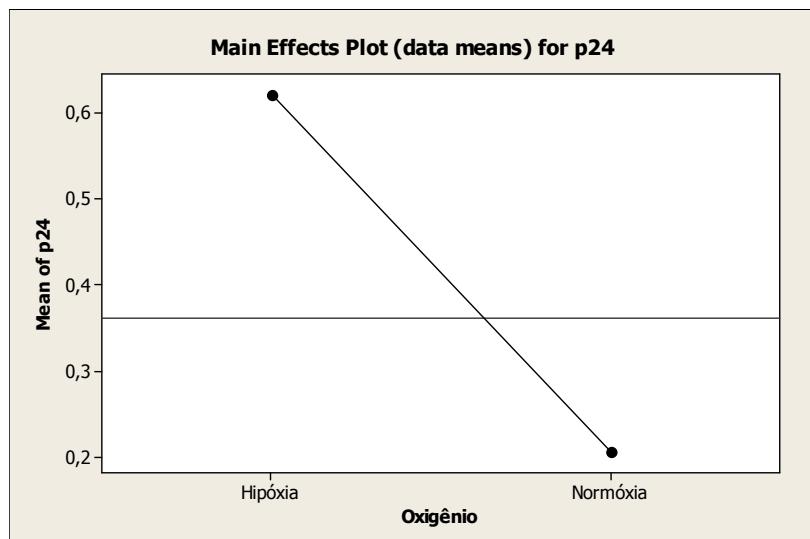


Gráfico 3: Perfis de Médias para a variável transformada de P6.

Nota-se que os resultados para estes dois metabólitos são diferentes, apesar de suas médias serem influenciados pelo mesmo fator, o Oxigênio, e, por isso, serem agrupados.

Para avaliarmos o efeito de interação entre os fatores, considere o metabólito P18. O Gráfico 4 apresenta as concentrações médias do metabólito P18 em combinações de Glicose e Oxigênio, mostrando que células submetidas inicialmente à condição Hipóxia possuem, em média, menos deste metabólito para nível de Glicoses Alta ao nível controle, ao passo que essa concentração média se altera se analisamos a condição Normóxia, sendo em média no nível de Glicose Alta maior que a média de Hipóxia no nível controle.

Tendo essas diferenciações entre metabólitos, conseguiu-se construir Grupos de Metabólitos com características que distinguem os mesmos fatores. Deste modo, pela ANOVA tem-se 10 grupos de metabólitos. A composição dos grupos formados pode ser encontrada no Apêndice 3. Com todos esses grupos, tem-se então uma seleção de quais seriam os principais metabólitos para se analisar juntos, com o objetivo de classificar uma nova célula nos níveis dos fatores Linhagem, Oxigênio e Glicose. Por exemplo, se a meta fosse predizer se uma nova célula veio da linhagem G2 ou da linhagem N3, ao invés de utilizar os 32 metabólitos para classificação, tomar-se-ia as medidas somente daqueles Metabólitos que, de alguma maneira (em termos de efeitos principais ou de interação), discriminam o fator Linhagem, ou seja, grupos 3, 5, 7, 8, 9 e 10.

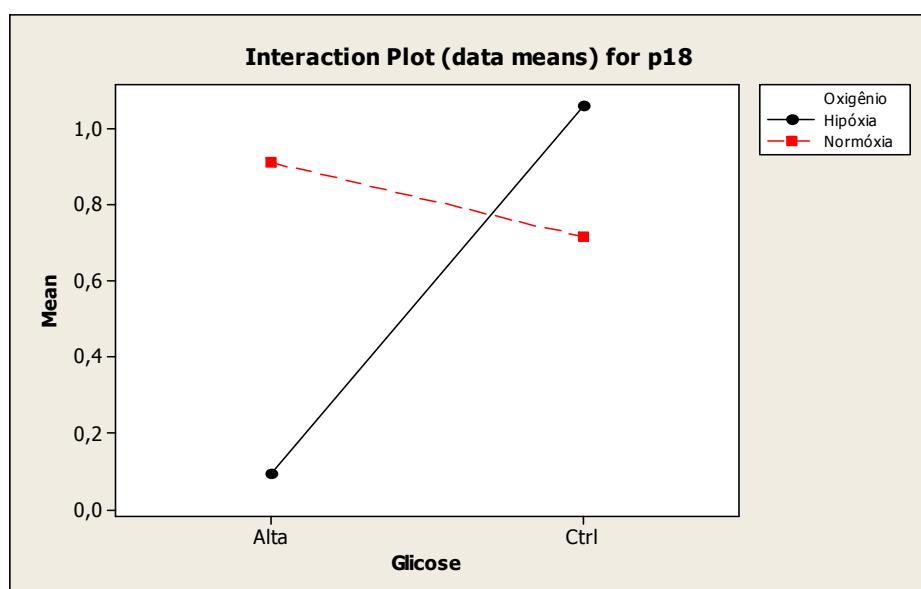


Gráfico 4 Perfis de Médias para a variável transformada de P18.

Seguindo esta abordagem, utilizar-se-ão técnicas estatísticas de segmentação, com base nos Grupos obtidos da ANOVA, aumentando assim a precisão de classificação de uma nova amostra.

5.2 Análise Discriminante com os Grupos de Metabólitos

A Análise Discriminante (ver, Johnson e Whichern, 2007) tem o objetivo de obter funções das variáveis (metabólitos) que segmentem as observações amostrais em grupos pré-especificados. Desta maneira, nosso objetivo em utilizar esta análise é encontrar Funções Discriminantes para validar o poder de segmentação dos Grupos de Metabólitos formados pela Análise de Variância, e verificar o quanto acurada é a separação das células do melanoma entre os níveis das condições experimentais Linhagem, Oxigênio e Glicose. Para avaliar a eficiência das funções discriminantes obtidas, deve-se olhar a porcentagem de acerto no critério de validação cruzada (*Cross Validation*).

Os níveis a serem discriminados são:

- Linhagem – G2 e N3
- Oxigênio – Normoxia e Hipóxia
- Glicose – Alta e Controle

Os resultados encontrados estão na Tabela 4.1 do Apêndice 4.

Os coeficientes dos metabólitos que compuseram as funções discriminantes para cada fator podem ser encontrados no Apêndice 4 (Tabelas 4.2 a 4.4).

As evidências obtidas mostram que os grupos de metabólitos da ANOVA são consistentes em explicar a variação entre os efeitos dos fatores sob estudo,

pois foi encontrada uma alta taxa de classificação, 72%, 78% e 84%, para os fatores Glicose, Oxigênio e Linhagem, respectivamente. Ainda, não foi necessário um grande número de metabólitos em cada função Discriminante, uma vez que a ANOVA facilitou a identificação e seleção dos Metabólitos importantes. Portanto, encontrou-se com esta seleção um bom critério de classificação de novas observações, respondendo, por um lado, ao objetivo do estudo.

Por exemplo, se uma nova célula têm os valores :

$$X_0 = (X_{01}, X_{02}, X_{03}, \dots, X_{031}, X_{032}),$$

$$\text{Seja } X_{03} = -0.16938, X_{05} = 1.459401, X_{016} = -1.39697$$

Desta maneira, calculamos o valor da função discriminante para os dois níveis do fator Linhagem, G2 e N3(ver função 3 no Apêndice 4).

Função Discriminante para o nível G2 = 27,30

Função Discriminante para o nível N3 = 26,40

Seleciona-se o nível que tem o maior valor na função discriminante, neste caso a nova célula é classificada no nível G2 do Fator Linhagem.

5.3 Regressão Logística

O modelo de Regressão Logística (ver, Agresti, 1996) pode ser usado com o mesmo objetivo da Análise Discriminante, isto é, o de encontrar a melhor combinação linear de variáveis (explicativas), que maximize a verossimilhança das frequências observadas na amostra, de forma a fazer previsões e a classificar novos casos. A regressão Logística incorpora três componentes: Aleatório, Sistemático e Logito. De maneira sucinta, a um logito é ajustado um modelo linear função de variáveis explicativas de interesse. Considere o seguinte modelo de Regressão Logística :

$$\ln\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k$$

$$p = \frac{e^{(\beta_0 + \beta_1 x_1 + \dots + \beta_k x_k)}}{1 + e^{(\beta_0 + \beta_1 x_1 + \dots + \beta_k x_k)}}$$

5.3.1 Breve Descrição dos Componentes do Modelo Logístico

O Componente Aleatório identifica uma variável resposta Y com distribuição de probabilidades Bernoulli (Sim=1 ou Não=0, por exemplo), em que p é a probabilidade da resposta Sim. Como exemplo, considere $Y=1$ se a amostra é do grupo Nomoxia e $Y=0$ se é do grupo Hipoxia.

O Componente Sistemático especifica as k variáveis explicativas (X 's), que são introduzidas no modelo de forma linear. O modelo de Regressão Logística assume uma relação linear entre as variáveis explicativas contínuas e a transformação *Logito* da variável resposta.

O Componente de Ligação $g(p)$, neste caso, o logito, é uma função monótona e diferenciável, que descreve a relação funcional entre p e as variáveis explicativas. O valor esperado p , neste caso, será a probabilidade de uma amostra pertencer ao nível de interesse do Fator em análise, isto é, $P(Y=1) = p$.

Com base nos Grupos de Metabólitos formados por intermédio das análises anteriores, serão ajustados três modelos logísticos, um para cada fator de controle do Estudo. Sendo assim, escolher-se-á um nível de interesse, para cada fator, e ajustar-se-ão os modelos. Após o ajuste, será possível obter a probabilidade (p) de uma determinada unidade amostral pertencer ao nível de interesse do fator em análise. Desta maneira, torna-se possível classificar as observações conforme as condições experimentais, atingindo o principal objetivo deste estudo.

A verificação da qualidade de Ajuste do modelo de regressão logística pode ser feita via Curva Roc, que é uma representação gráfica da sensibilidade,

proporção de acertos do modelo, versus (1-especificidade), proporção de classificações errôneas, para uma variável binária Y . Esta curva pode ser interpretada como a relação de verdadeiros positivos ($V+$) versus a proporção de falsos positivos($F+$) e, também, ser vista como uma medida de prognóstico, auxiliando no ajuste e na seleção de modelos. A seguinte matriz de contingência descreve estas quantidades:

Matriz de Contingência

		Resposta Preditada pelo Modelo	
		$Y=1$	$Y=0$
Resposta Observada	$Y=1$	$V+$	$F-$
$Y=0$		$F+$	$V-$

em que:

- **$V+$** : Verdadeiro Positivo (observações preditas como verdadeiras pelo modelo quando de fato essas observações são verdadeiras)
- **$V-$** : Verdadeiro Negativo (observações preditas como Falsas pelo modelo quando de fato essas observações são faltas)
- **$F-$** : Falso Negativo (observações preditas como falsas pelo modelo quando de fato elas são verdadeiras)
- **$F+$** : Falso Positivo (observações preditas como verdadeiras pelo modelo quando de fato elas são falsas)

Destes modos, tem-se as definições:

- **Sensibilidade** : $\frac{V+}{(V+)+(F-)} = P(Y=1 \text{ modelo} | Y=1 \text{ observado})$
- **Especificidade** : $\frac{V-}{(V-)+(F-)} = P(Y=0 \text{ modelo} | Y=0 \text{ observado})$
- **(1-Especificidade)** : $\frac{F-}{(V-)+(F-)}$
- **Valor Preditivo** : $\frac{(V+)+(V-)}{n}$: proporção geral de acertos

Com o auxílio da Curva Roc torna-se possível averiguar o ajuste de modelos em Regressões Logísticas. Todavia a área sobre a curva representa a quantidade de acertos e erros, relativamente ao nível de interesse do Fator em estudo. Sendo assim, quanto maior a sensibilidade e a especificidade, maior será a área sob a curva, que será uma medida de ajuste e pode ser usada como o poder preditivo do modelo. A reta $Y = X$ representa a aleatoriedade, ou seja, o modelo com área sobre a curva Roc de 50% é equivalente a tomar-se decisão jogando uma moeda e definindo um dos lados como de interesse. Um modelo com bom ajuste terá área sob a curva Roc maior que 50% e, quando possível, próxima de um.

Tabela 3: Resultados das Regressões Logísticas

Fator	Taxa de Acerto
<i>Linhagem</i>	87,50%
<i>Oxigênio</i>	84,30%
<i>Glicose</i>	81,20%

A validação das suposições e resultados inerentes a esta análise podem ser encontradas no Apêndice 5.

Os três modelos Logísticos apresentaram resultados satisfatórios pela análise da Curva Roc (Gráficos 5.1, 5.2 e 5.3). As áreas sob a curva variaram de 85% a 94%, indicando que os modelos obtiveram boas classificações em relação aos níveis de interesse de cada fator. Pelo Teste da Qualidade de Ajuste de Pearson, em que a hipótese nula é H_0 : Modelo está bem ajustado, obtiveram-se resultados em que os modelos para os fatores Linhagem, Oxigênio e Glicose estão bem ajustados (valores-p: 0,3677, 0,9305 e 0,3355, respectivamente). Com base no modelo ajustado, por exemplo, a tabela 5.1, do Apêndice 5, indica que a probabilidade da amostra X_0 pertencer ao nível G2 é calculada da seguinte maneira :

$$p = \frac{e^{(-8.55 - 6.28x_3 + 8.02x_5 + 1.98x_{16})}}{1 + e^{(-8.55 - 6.28x_3 + 8.02x_5 + 1.98x_{16})}} = 0.809$$

Em que, $X_0 = (X_{01}, X_{02}, X_{03}, \dots, X_{031}, X_{032})$,

e $X_{03} = -0.16938$, $X_{05} = 1.459401$, $X_{016} = -1.39697$

Desta forma, assumindo um corte de 50%, a amostra X_0 é classificada no nível G2 do Fator Linhagem. Note que o resultado obtido por intermédio da Regressão Logística corrobora com o obtido na Análise Discriminante.

5.4 Conclusões

Identificou-se neste estudo, via ANOVA, que existem variáveis que diferenciam as células do melanoma com relação às condições experimentais de Linhagem, Oxigênio e Glicose. A análise de Componentes Principais não se mostrou útil para discriminar os grupos quando todos os metabólitos são analisados conjuntamente. Além disso, também se obteve critérios probabilísticos e classificatórios, via Regressão Logística e Análise Discriminante, respectivamente, para mapear a origem de uma nova célula de melanoma conhecendo seu perfil metabólico e observando principalmente os metabólitos que foram apontados neste estudo como mais importantes para segmentação dos níveis de Linhagem, Oxigênio e Glicose.

Apêndice

Apêndice 1 : Tabelas e Gráficos Descritivos

Tabela 1.1: Tabela Descritiva com os Dados Originais

Variable	Mean	Std Dev	Minimum	Maximum	Coeff of Variation
P1	0.0057478	0.0019638	0.00369	0.01225	34.166366
P2	0.0061738	0.0024983	0.003348	0.01417	40.4656101
P3	0.0160427	0.006367	0.008642	0.03989	39.6880403
P4	0.0114174	0.0049296	0.006376	0.02461	43.1763315
P5	0.061625	0.0147381	0.02949	0.08649	23.9158411
P6	0.012497	0.0028942	0.007412	0.0182	23.1595309
P7	0.006349	0.0025692	0.003422	0.01473	40.4667579
P8	0.02956	0.0244643	0.006073	0.1027	82.7615714
P9	0.0143557	0.0043013	0.009601	0.02656	29.9623183
P10	0.0130234	0.0058386	0.007306	0.02973	44.8318284
P11	0.0109457	0.0027025	0.004717	0.01718	24.689745
P12	0.0033448	0.00081962	0.002074	0.004888	24.5039841
P13	0.0246966	0.010981	0.01037	0.06465	44.4636834
P14	0.0057586	0.0033319	0.002131	0.01366	57.8588494
P15	0.0064434	0.0061308	0.001503	0.02835	95.1476762
P16	0.0122243	0.011808	0.001396	0.04387	96.5945367
P17	0.0053327	0.0023248	0.00224	0.01345	43.5962198
P18	0.0295534	0.0144939	0.00543	0.05565	49.0432765
P19	0.0239799	0.0089794	0.003876	0.04378	37.4453008
P20	0.0079514	0.0022924	0.00349	0.01494	28.8302151
P21	0.0513384	0.0177891	0.02209	0.115	34.6506557
P22	0.0147561	0.0036837	0.007347	0.02435	24.9638005
P23	0.004936	0.0017034	0.002794	0.008892	34.5098591
P24	0.0198325	0.0110368	0.008877	0.05693	55.6501545
P25	0.0041612	0.0012695	0.002062	0.008758	30.5087878

P26	0.0121254	0.0046158	0.003749	0.02057	38.0669006
P27	0.0103496	0.0045899	0.002355	0.019	44.3485255
P28*	0.0233879	0.0458072	0.006556	0.2594	195.858674
P29	0.0579728	0.0147007	0.03011	0.08492	25.357878
P30	0.0027824	0.000793742	0.001097	0.004272	28.5271877
P31	0.0235193	0.0083509	0.008086	0.0374	35.5064819
Poutros	0.4678159	0.0414262	0.354087	0.535352	8.8552345

Metabólito Modificado

Variabel	Mean	Std Dev	Minimum	Maximum	Coef. Of Variation
P28	0.0158	0.0156	0.0066	0.0877	98.923

Tabela 1.2 : Matriz(Parcial) de Correlação

p1	p2	p3	p4	p5	p6	p7	p8	p9	p10	p11	p12	p13	p14	p15	p16	p17	p18	p19	p20	p21	p22	p23	p24
1.0000	0.8427	0.3511	0.8983	-0.3190	0.0749	0.6546	0.4348	0.8919	0.8618	-0.2592	-0.5040	-0.1923	0.1446	-0.4785	0.3835	-0.2474	-0.5786	-0.6223	-0.4922	-0.4176	-0.6706	-0.5868	-0.3927
0.8427	1.0000	0.5180	0.9059	-0.1894	0.0360	0.6820	0.4646	0.7330	0.8748	-0.2891	-0.5078	-0.2645	0.0015	-0.4795	0.2960	-0.2046	-0.5046	-0.4952	-0.5100	-0.4427	-0.6031	-0.6008	-0.3415
0.3511	0.5180	1.0000	0.3920	0.2852	0.0500	0.4165	-0.0288	0.2324	0.4082	-0.0999	-0.2077	-0.1703	-0.2452	-0.1627	-0.2162	-0.1902	-0.1934	-0.2789	-0.2269	-0.2203	-0.2167	-0.3875	-0.0575
0.8983	0.9059	0.3920	1.0000	-0.3625	0.0545	0.7731	0.5731	0.6868	0.9587	-0.3474	-0.5954	-0.1553	0.3289	-0.4492	0.5156	-0.3661	-0.7267	-0.7131	-0.4544	-0.3788	-0.6886	-0.5209	-0.3446
-0.3190	-0.1894	0.2852	-0.3625	1.0000	-0.2548	-0.2926	-0.3675	-0.5061	-0.3911	-0.0682	0.0887	-0.0044	-0.4045	0.3596	-0.5482	0.3691	0.3493	0.3645	0.0632	0.0988	0.2237	0.2811	0.3129
0.0749	0.0360	0.0500	0.0545	-0.2548	1.0000	0.1973	0.2394	0.3263	0.0874	0.7116	0.5348	-0.4112	0.1131	-0.5714	0.2570	-0.1581	0.1297	-0.0223	-0.2824	-0.4813	-0.3759	0.0712	-0.6447
0.6546	0.6820	0.4165	0.7731	-0.2926	0.1973	1.0000	0.2438	0.6911	0.8059	-0.1764	-0.3502	-0.0159	0.3616	-0.3279	0.2717	-0.3137	-0.6510	-0.7706	-0.3146	-0.2312	-0.0796	-0.5131	-0.4103
0.4348	0.4646	-0.0288	0.5731	-0.3675	0.2394	0.2438	1.0000	0.4967	0.4838	-0.0258	-0.3924	-0.4238	0.3304	-0.4705	0.8872	-0.3314	-0.4434	-0.1745	-0.3201	-0.5149	-0.4796	-0.0003	-0.3887
0.8919	0.7330	0.2324	0.8688	-0.5061	0.3263	0.6911	0.4967	1.0000	0.8808	-0.0055	-0.3156	-0.3102	0.3148	-0.5446	0.5092	-0.3503	-0.5842	-0.6501	-0.4985	-0.5200	-0.7120	-0.4560	-0.5050
0.8618	0.8748	0.4082	0.9587	-0.3911	0.0874	0.8059	0.4838	0.8809	1.0000	-0.2916	-0.5077	-0.1681	0.2551	-0.4250	0.4116	-0.3782	-0.7056	-0.7183	-0.4600	-0.4134	-0.6584	-0.5621	-0.3370
-0.2592	-0.2891	-0.0999	-0.3474	-0.0862	0.7116	-0.1764	-0.0258	-0.0055	-0.2916	1.0000	0.7907	-0.5816	-0.2117	-0.4390	-0.0832	0.3203	0.5720	0.4222	-0.2408	-0.4944	-0.1779	0.2392	-0.5090
-0.5040	-0.5078	-0.2077	-0.5054	0.0897	0.5348	-0.3502	-0.3924	-0.3156	-0.5077	0.7907	1.0000	-0.3300	-0.2580	-0.1283	-0.3876	0.3996	0.6095	0.4655	-0.0940	-0.2266	0.0825	0.4018	-0.2148
-0.1923	-0.2645	-0.1703	-0.1553	-0.0044	-0.4112	-0.1519	-0.4238	-0.3102	-0.1881	-0.5816	-0.3300	1.0000	0.3673	0.4152	-0.2035	-0.2070	-0.1680	-0.2368	0.6723	0.9508	0.4818	-0.1345	0.3262
0.1446	0.0015	-0.2452	0.3289	-0.4045	0.1131	0.3616	0.3304	0.3148	0.2551	-0.2117	-0.2585	0.3673	1.0000	-0.0060	0.8089	-0.4971	-0.5815	-0.6225	0.1517	0.2297	-0.2038	0.1810	-0.1175
-0.4785	-0.4795	-0.1627	-0.4492	0.3596	-0.5714	-0.3279	-0.4705	-0.5446	-0.4250	-0.4390	-0.1283	0.4152	-0.0060	1.0000	-0.2725	-0.0623	-0.252	0.0157	0.4994	0.5578	0.5871	0.3393	0.8245
0.3835	0.2960	-0.2162	0.5156	-0.5482	0.2570	0.2717	0.8872	0.5092	0.4116	-0.0832	-0.3876	-0.2035	0.6069	-0.2725	1.0000	-0.5075	-0.5969	-0.3731	-0.1447	-0.3070	-0.3759	0.0937	-0.2817
-0.2474	-0.2046	-0.1902	-0.3661	0.3891	-0.1581	-0.3137	-0.3314	-0.3503	-0.3782	0.3203	0.3996	-0.2070	-0.4971	-0.0823	-0.5075	1.0000	0.6088	0.5683	-0.2714	-0.1129	-0.1125	0.0574	-0.0507
-0.5786	-0.5046	-0.1934	-0.7267	0.3493	0.1297	-0.6510	-0.4434	-0.5842	-0.7056	0.5720	0.6995	-0.1680	-0.5815	-0.0252	-0.5969	0.6605	1.0000	0.0272	0.0114	0.0051	0.2529	0.2535	-0.0654
-0.8223	-0.4952	-0.2789	-0.7131	0.3645	-0.0223	-0.7706	-0.1745	-0.6561	-0.7183	0.4222	0.4655	-0.2368	-0.8225	0.0157	-0.3731	0.5683	0.8272	1.0000	0.0428	-0.0251	0.3653	0.4024	0.1111
-0.4922	-0.5100	-0.2289	-0.4544	0.0632	-0.2824	-0.3146	-0.3201	-0.4985	-0.4600	-0.2408	-0.0940	0.6723	0.1517	0.4994	-0.1447	-0.2714	0.0114	0.0428	1.0000	0.7916	0.8339	0.0140	0.3220
-0.4176	-0.4427	-0.2203	-0.3788	0.0988	-0.4813	-0.2312	-0.5149	-0.5200	-0.4134	-0.4944	-0.2266	0.9508	0.2297	0.5578	-0.3070	-0.1129	0.0051	-0.0251	0.7916	1.0000	0.6780	0.0010	0.4596
-0.6706	-0.6031	-0.2167	-0.6886	0.2237	-0.3759	-0.6079	-0.4798	-0.7120	-0.6584	-0.1779	0.0825	0.4818	-0.2038	0.5871	-0.3759	-0.1125	0.2529	0.3653	0.8339	0.6780	1.0000	0.1994	0.5353
-0.5868	-0.6008	-0.3875	-0.5209	0.2811	0.0712	-0.5131	-0.0003	-0.4566	-0.5621	0.2392	0.4018	-0.1348	0.1810	0.3393	0.0937	0.0574	0.2353	0.4024	0.0140	0.0010	0.1994	1.0000	0.3591
-0.3927	-0.3415	-0.0575	-0.3446	0.3129	-0.6477	-0.4103	-0.3887	-0.5050	-0.3370	-0.5090	-0.2148	0.3202	-0.1175	0.8245	-0.2817	-0.0507	-0.0654	0.1111	0.3220	0.4596	0.5353	0.3591	1.0000
0.0851	0.0822	0.2830	-0.0362	0.1421	-0.3388	0.0899	-0.6288	-0.0796	0.0895	-0.1977	0.0877	0.2042	-0.3910	0.0629	-0.7162	0.2631	0.0919	-0.1084	0.0179	0.1770	0.1597	-0.4124	0.1875
-0.4832	-0.4586	-0.0179	-0.8023	0.3071	-0.1119	-0.3070	-0.7771	-0.4885	-0.4478	0.3327	0.5860	0.0088	-0.5110	0.1005	-0.8361	0.5304	0.6385	0.4142	-0.0098	0.1118	0.2442	0.0971	0.2110
-0.5539	-0.3972	-0.0288	-0.5955	0.3428	0.1330	-0.4229	-0.3021	-0.4980	-0.5277	0.6314	0.6474	-0.4459	-0.8018	-0.0579	-0.4910	0.6429	0.7848	0.7497	-0.2526	-0.2707	0.0437	0.3063	-0.0575
0.3710	0.2709	0.1440	0.3487	-0.0439	-0.4620	0.2454	-0.0334	0.1948	0.3605	-0.7662	-0.6067	0.0315	0.1074	0.2977	0.0587	-0.4582	-0.6306	-0.4953	0.0916	0.2146	0.1113	-0.3206	0.4141
-0.4482	-0.4468	-0.2558	-0.6364	0.1113	0.3373	-0.5513	-0.3421	-0.4133	-0.6162	0.6725	0.7524	-0.2148	-0.5187	-0.2340	-0.4374	0.4996	0.8266	0.6847	0.0982	-0.0733	0.3031	0.1196	-0.3033
-0.2169	-0.2798	0.1558	-0.3827	0.2946	-0.1160	-0.2457	-0.6739	-0.3400	-0.3574	0.0594	0.2077	0.3572	-0.3422	0.0292	-0.7508	0.3612	0.4756	0.2800	0.1656	0.3935	0.2560	-0.1385	0.1485
-0.1036	-0.1287	-0.1910	-0.3118	-0.0633	0.2573	-0.3268	-0.1545	-0.0505	-0.2544	0.5241	0.4438	-0.4066	-0.5631	-0.2674	-0.2283	0.3500	0.5184	0.4535	-0.1285	0.0992	-0.2093	-0.3498	
0.1381	0.0498	0.2096	0.0625	-0.1099	0.2507	0.4083	-0.4000	0.1484	0.1219	0.2508	0.3068	0.0743	0.0073	-0.0896	-0.2973	0.0605	-0.0424	-0.3827	0.0280	0.0185	-0.1100	-0.2913	-0.2658

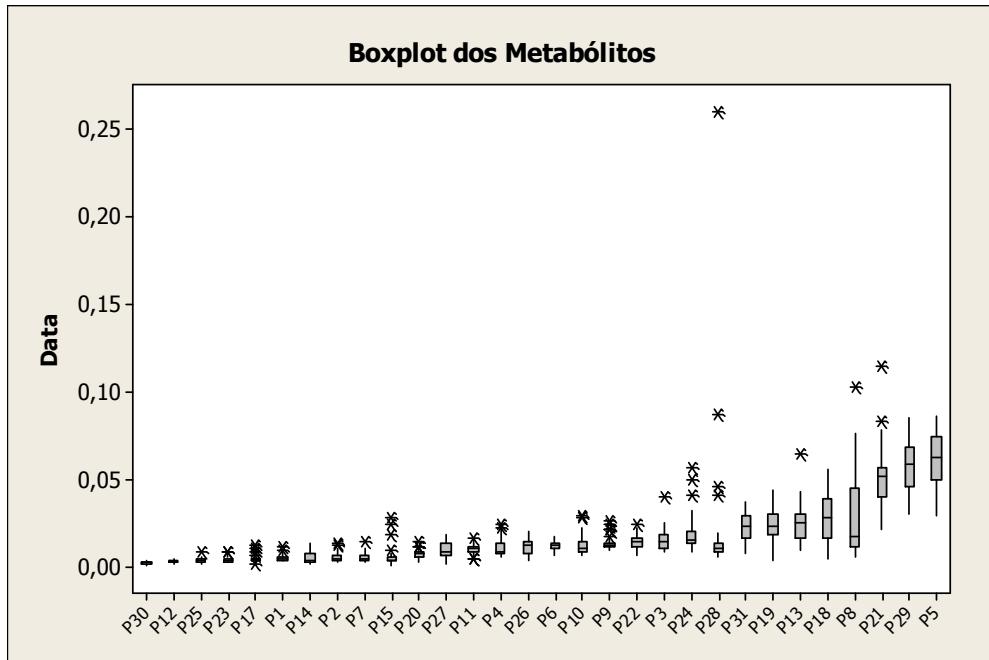


Gráfico 1.1: *Box-Plot dos Metabólitos*

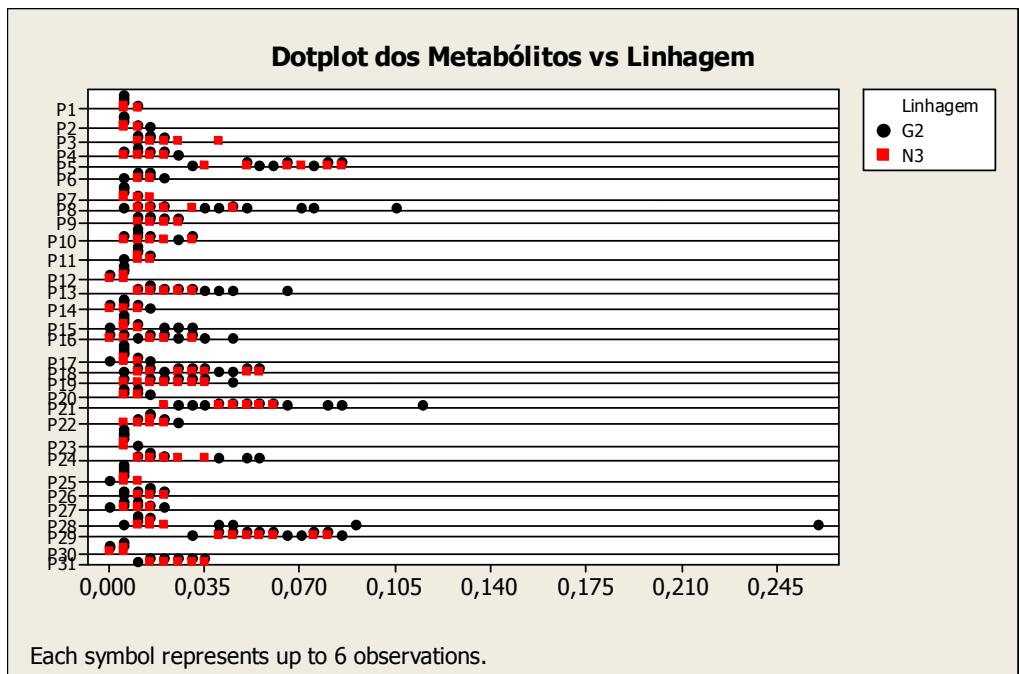


Gráfico 1.2: Dotplot dos Metabólitos para o Fator Linhagem

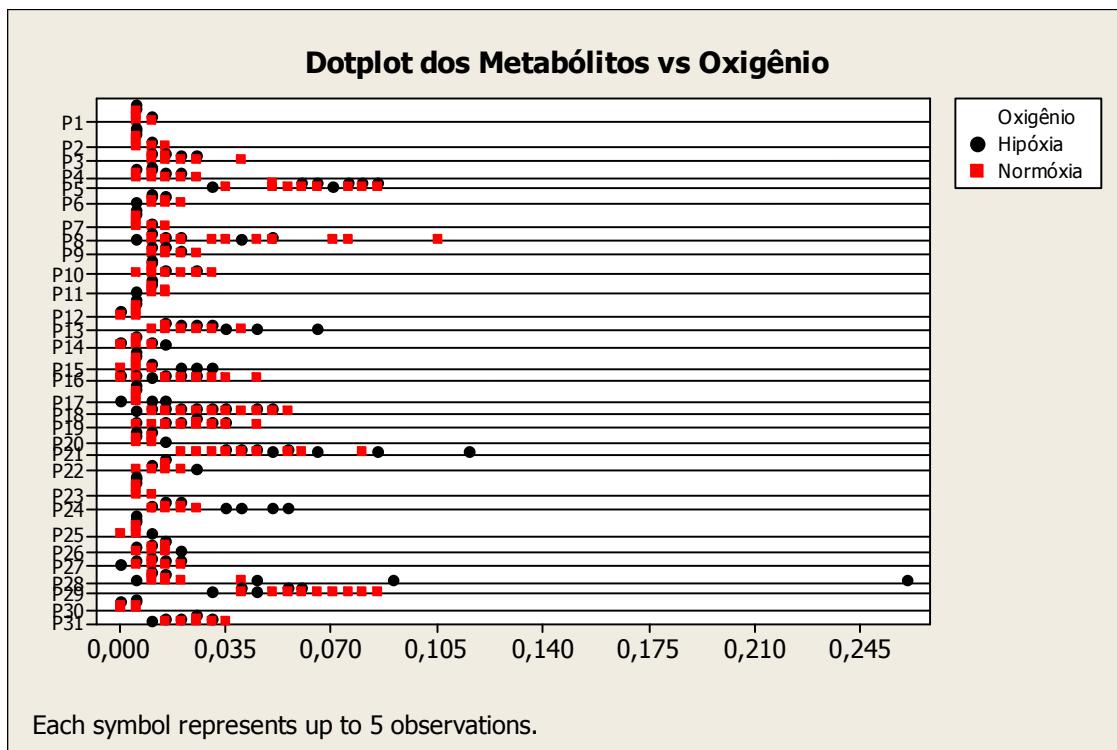


Gráfico 1.3: Dotplot dos Metabólitos para o Fator Oxigênio

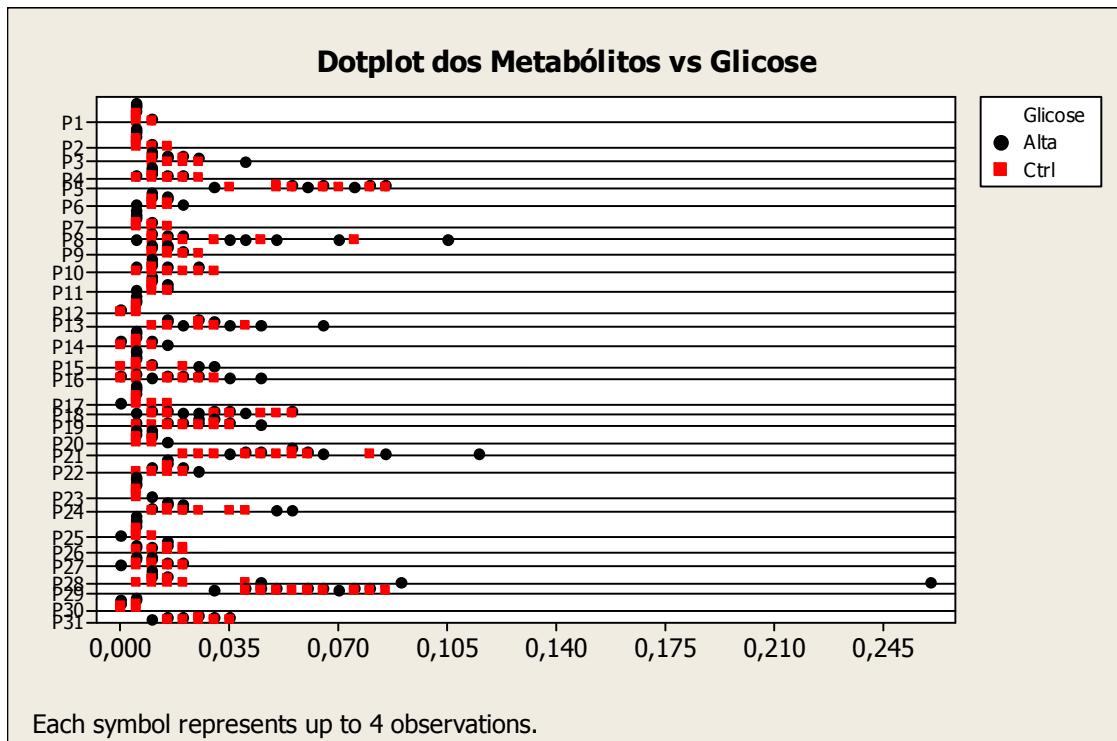


Gráfico 1.4: Dotplot dos Metabólitos para o Fator Glicose

Metabólito 28 – Modificado

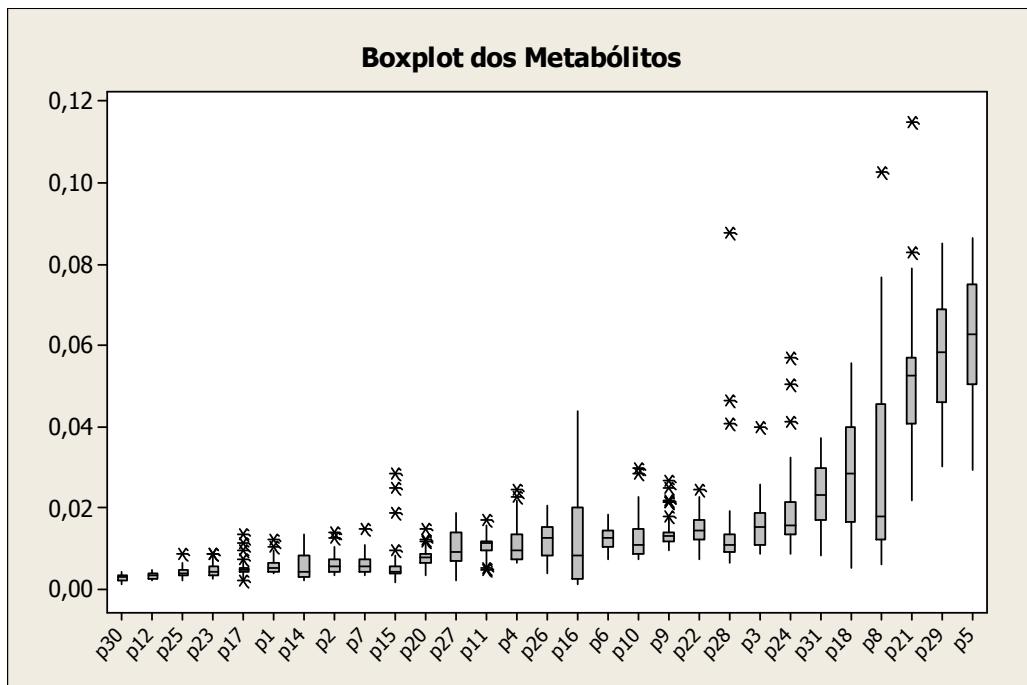


Gráfico 1.5: Box-Plot dos Metabólitos

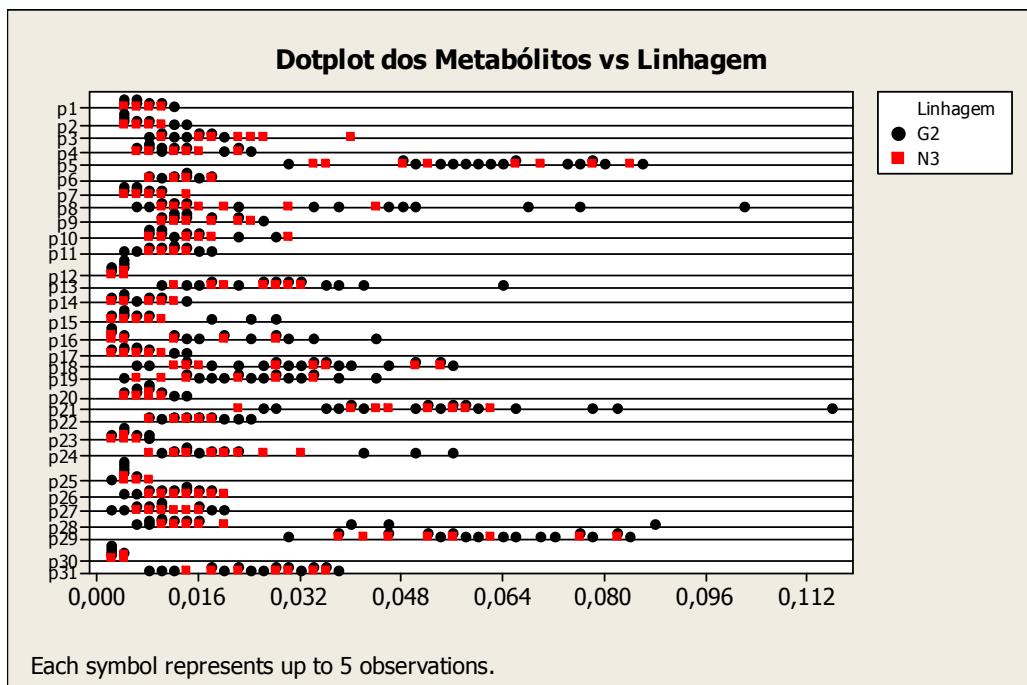


Gráfico 1.6: Dotplot dos Metabólitos para o Fator Linhagem

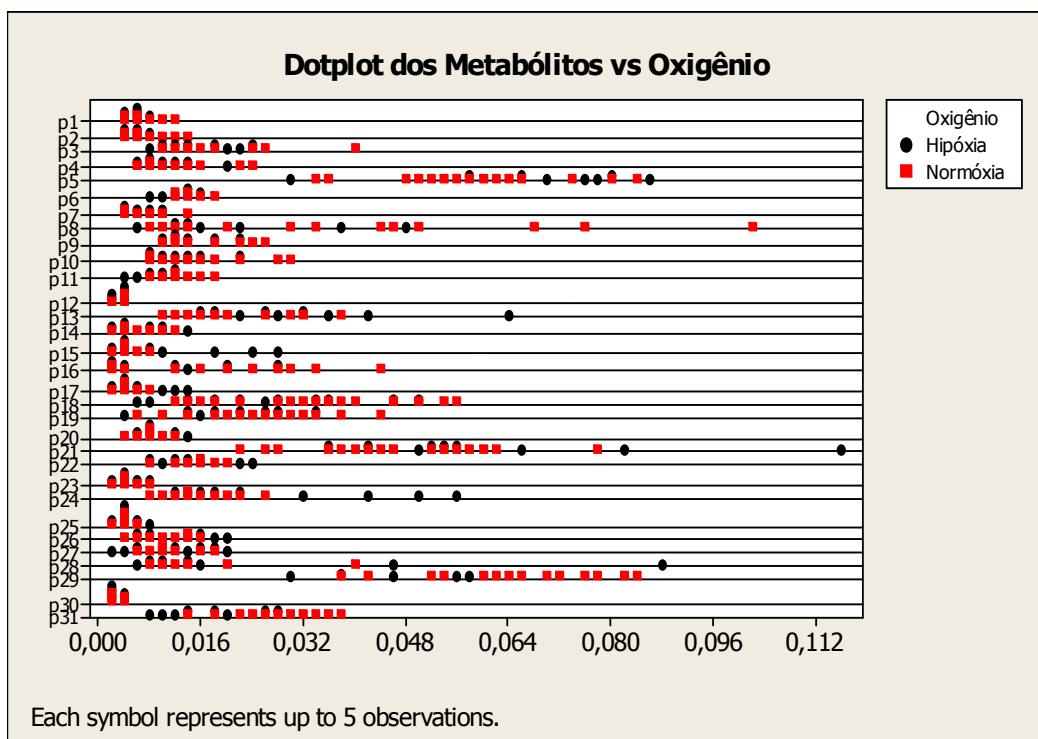


Gráfico 1.7: Dotplot dos Metabólitos para o Fator Oxigênio

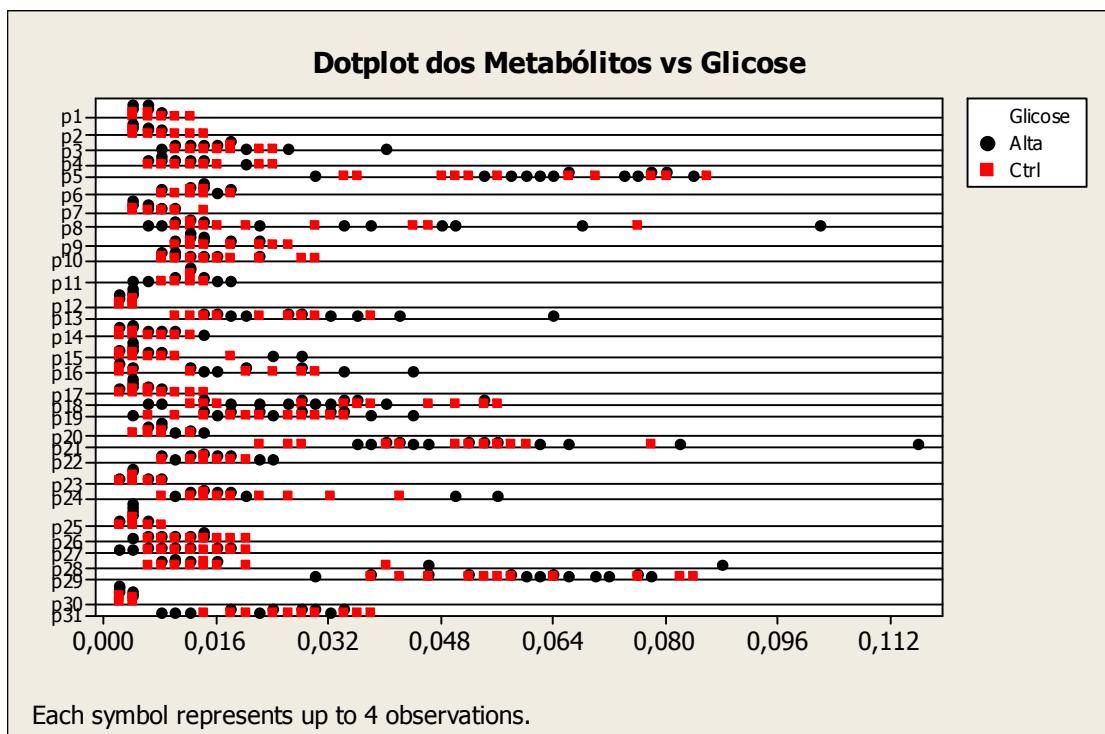


Gráfico 1.8: Dotplot dos Metabólitos para o Fator Glicose

Apêndice 2: Análise de Componentes Principais

Dados Transformados – Matriz de Correlação

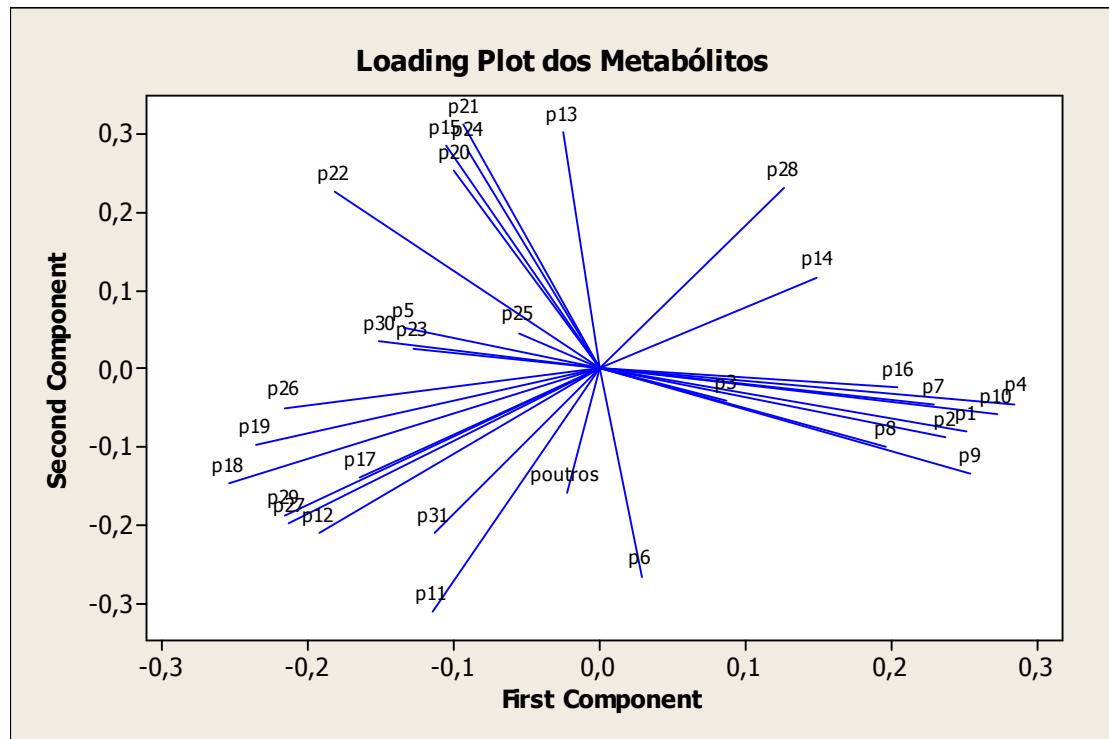


Gráfico 2.1: Loading Plot dos Componentes 1 e 2 (Gráfico de pesos)

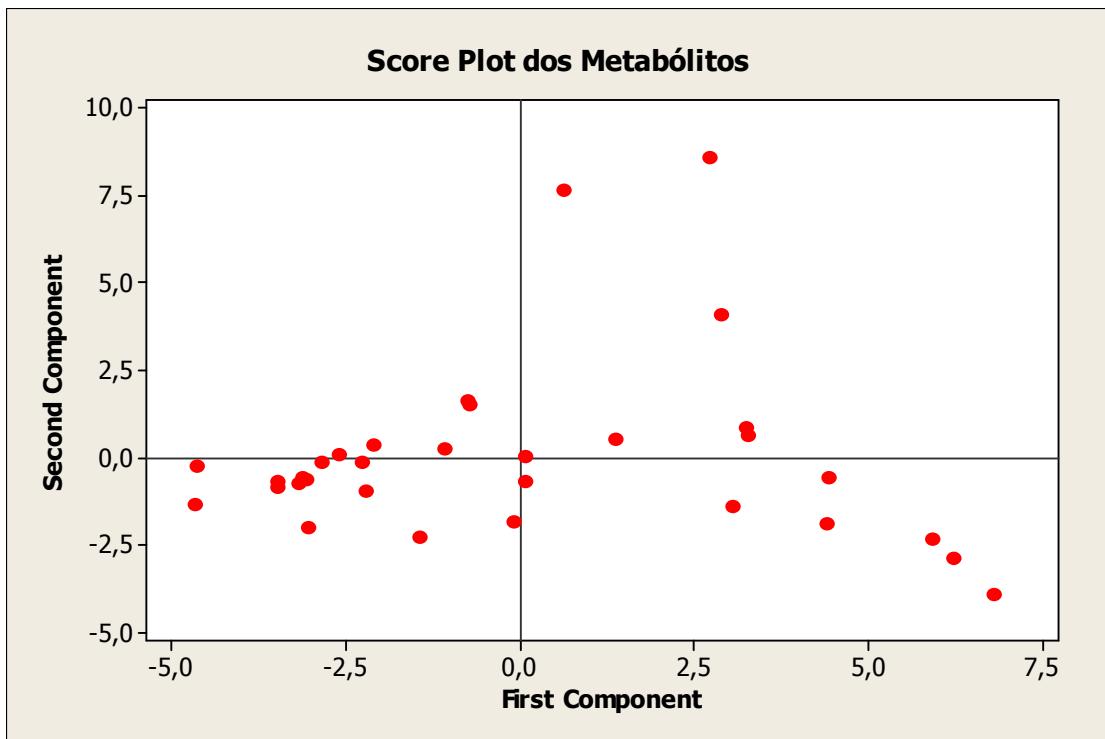


Gráfico 2.2: Gráfico de Dispersão: Componente1 vs Componente2

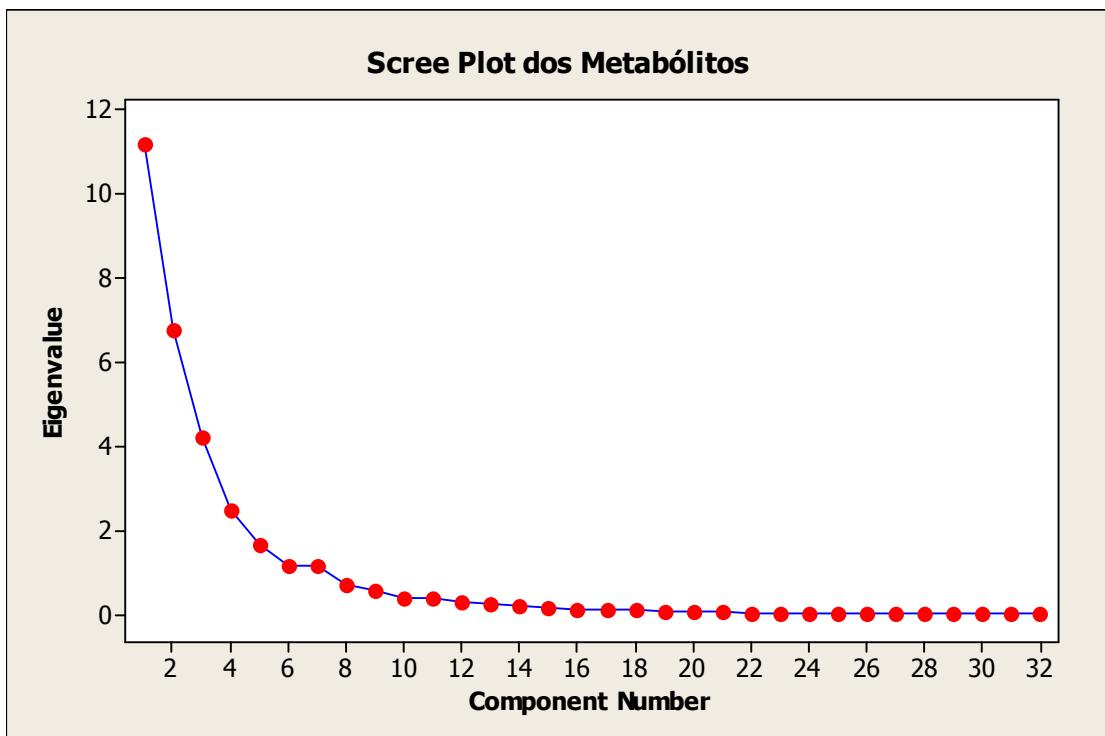


Gráfico 2.3: Scree Plot da Análise de Componentes Principais

Análise de Componentes Principais: Metabólitos

Matriz de Correlação:

Auto-valor: 11,196 6,745 4,207 2,462 1,642 1,164 1,138 0,707 0,547

Proporção: 0,350 0,211 0,131 0,077 0,051 0,036 0,036 0,022 0,017

Acumulado: 0,350 0,561 0,692 0,769 0,820 0,857 0,892 0,914 0,932

**Cargas das Componentes Principais:*

Variável	PC1	PC2	PC3	PC4
P1	0,252	-0,081	-0,156	-0,043
P2	0,237	-0,088	-0,171	-0,154
P3	0,086	-0,041	-0,274	-0,123
P4	0,285	-0,045	-0,093	-0,057
P5	-0,135	0,052	-0,111	-0,294
P6	0,028	-0,266	0,122	0,281
P7	0,229	-0,047	-0,166	0,141
P8	0,196	-0,100	0,277	-0,142
P9	0,254	-0,135	-0,057	0,069
P10	0,272	-0,058	-0,137	-0,039
P11	-0,115	-0,311	0,095	0,166
P12	-0,192	-0,210	0,036	0,171
P13	-0,025	0,303	-0,073	0,286
P14	0,149	0,117	0,217	0,289
P15	-0,106	0,285	0,024	-0,119
P16	0,204	-0,024	0,341	0,019
P17	-0,164	-0,139	-0,121	-0,178
P18	-0,254	-0,148	-0,019	-0,027
P19	-0,235	-0,097	0,104	-0,208
P20	-0,100	0,253	0,065	0,288

Variável	PC1	PC2	PC3	PC4
P21	-0,093	0,312	-0,038	0,241
P22	-0,182	0,226	0,038	0,094
P23	-0,128	0,026	0,314	-0,130
P24	-0,091	0,281	-0,009	-0,264
P25	-0,056	0,046	-0,417	0,030
P26	-0,217	-0,050	-0,230	-0
P27	-0,213	-0,198	-0,016	-0,173
P28	0,126	0,231	-0,142	-0,113
P29	-0,216	-0,189	-0,008	0,158
P30	-0,151	0,035	-0,276	0,123
P31	-0,113	-0,211	-0,039	0,024
Poutros	-0,023	-0,159	-0,270	0,334

Gráficos de Dispersão das Componentes

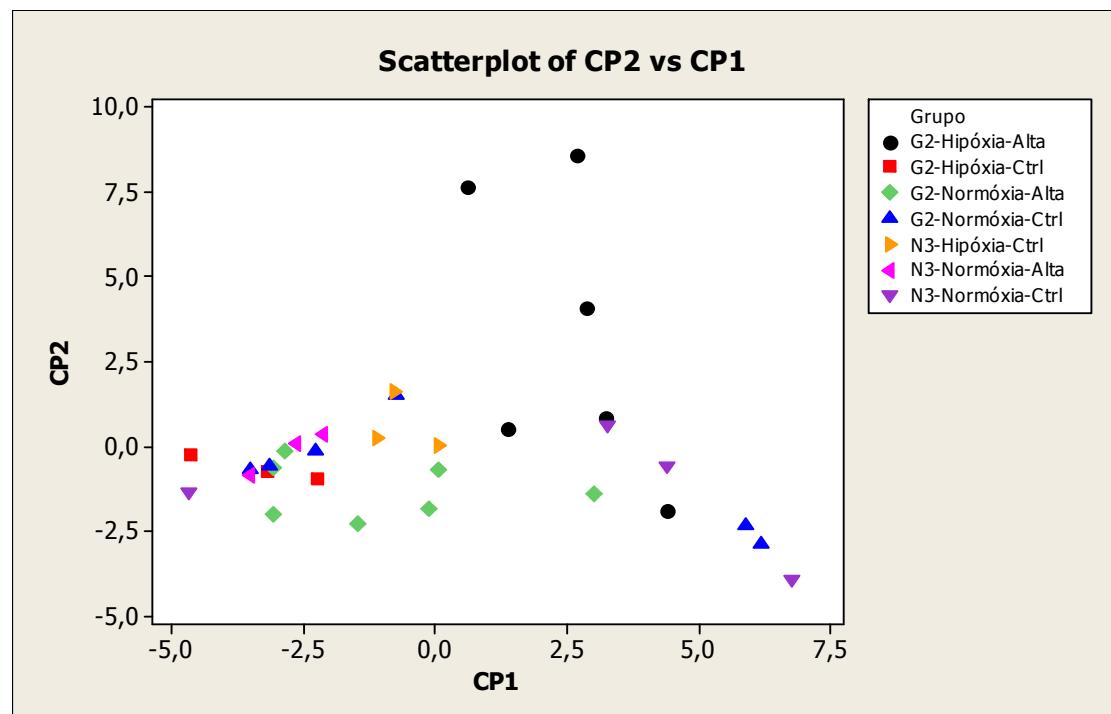


Gráfico 2.4: Dispersão: Componente 2 vs Componente1

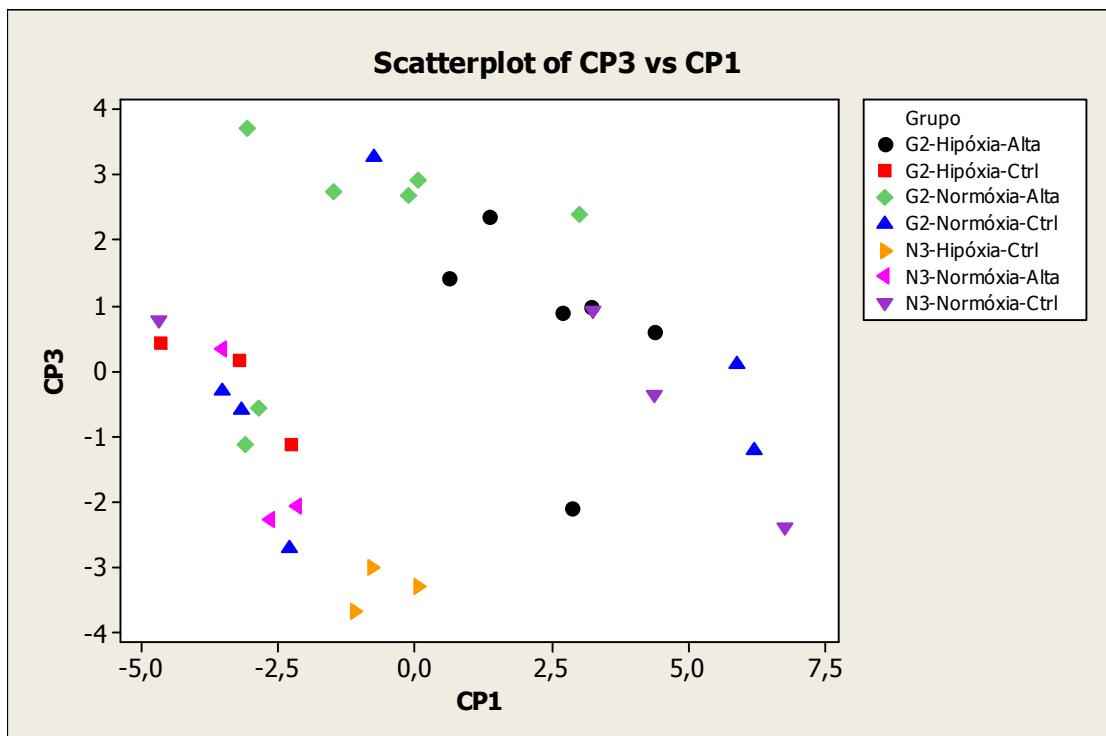
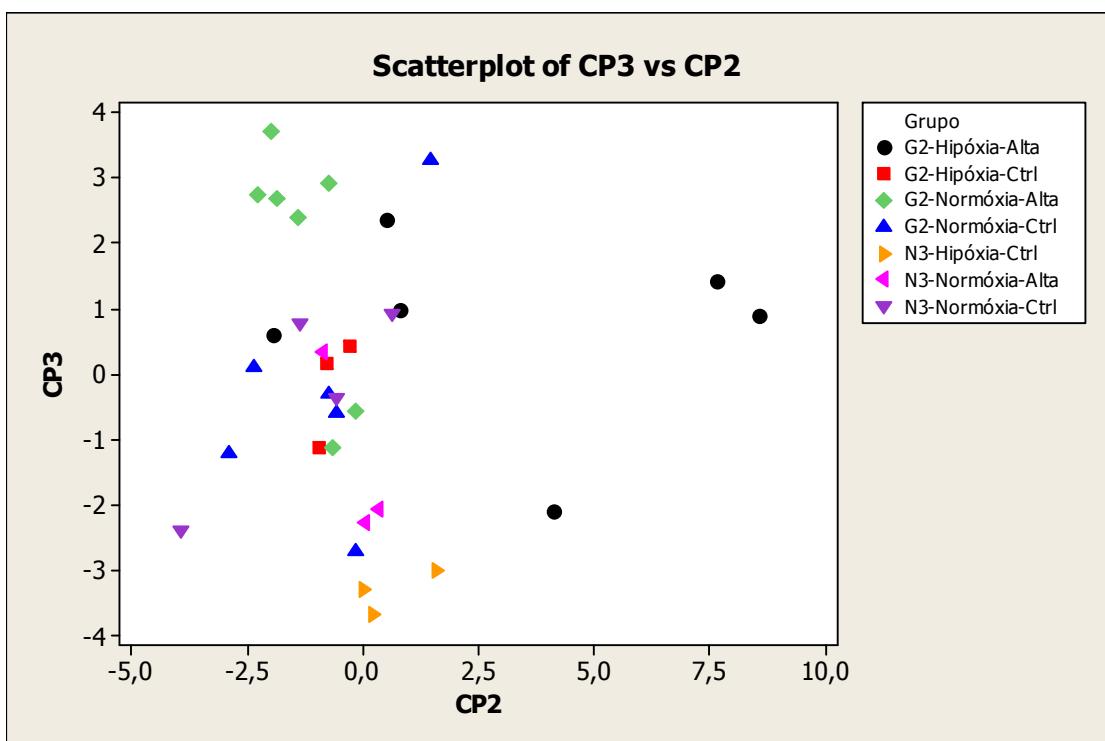
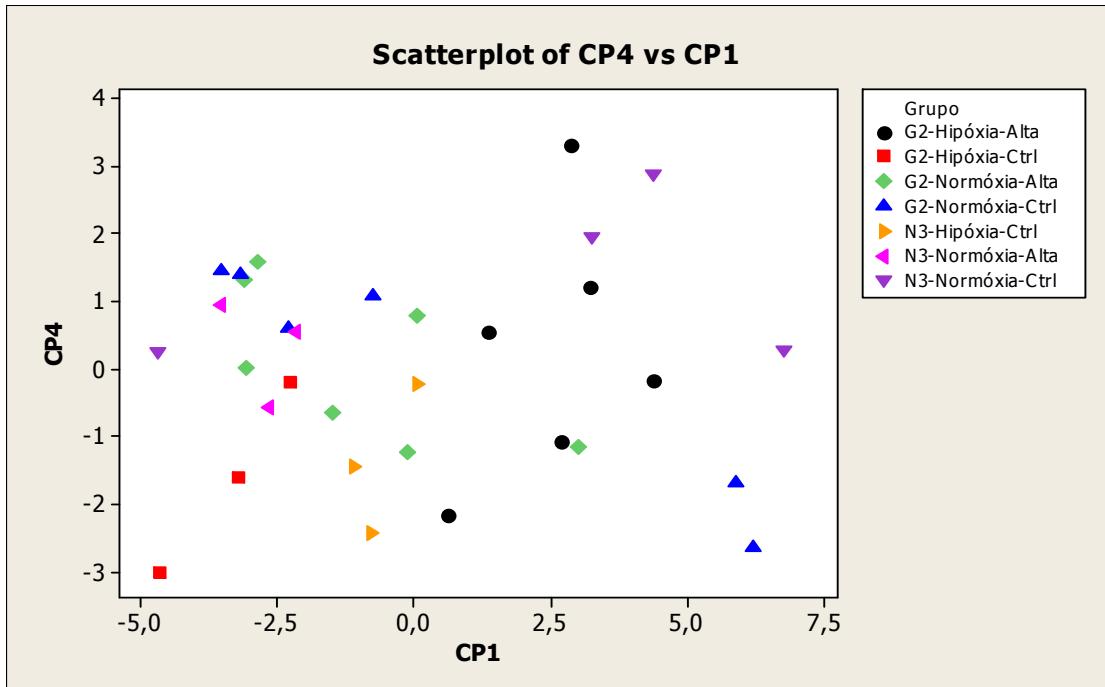
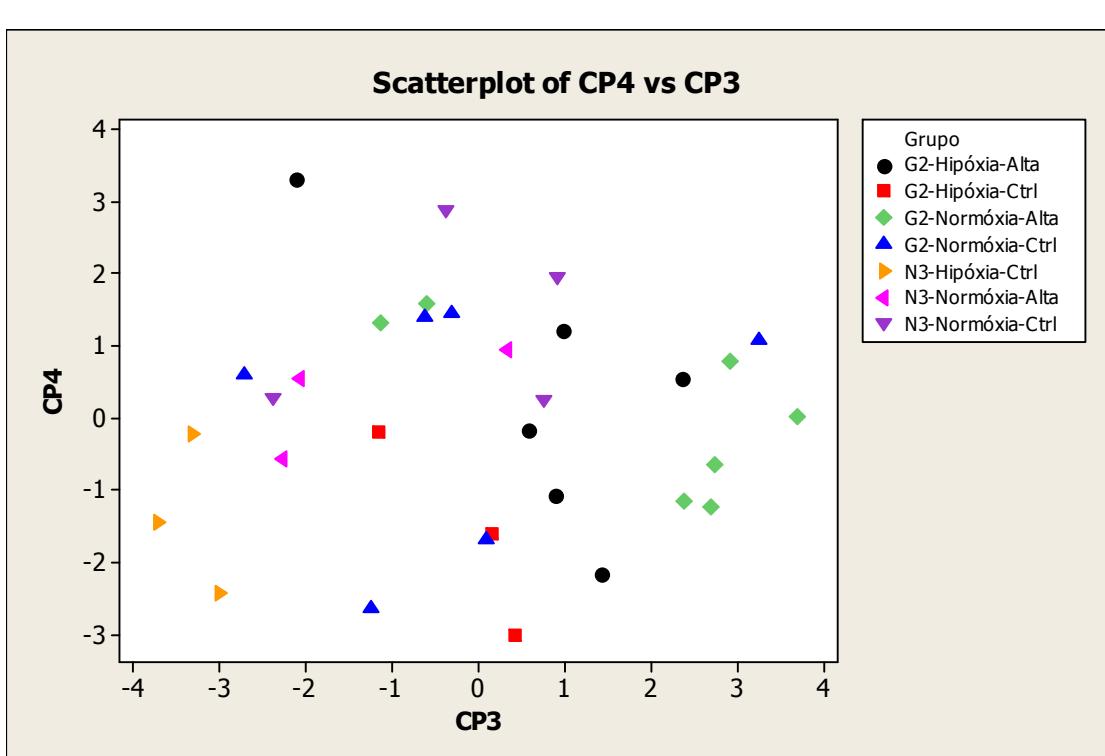
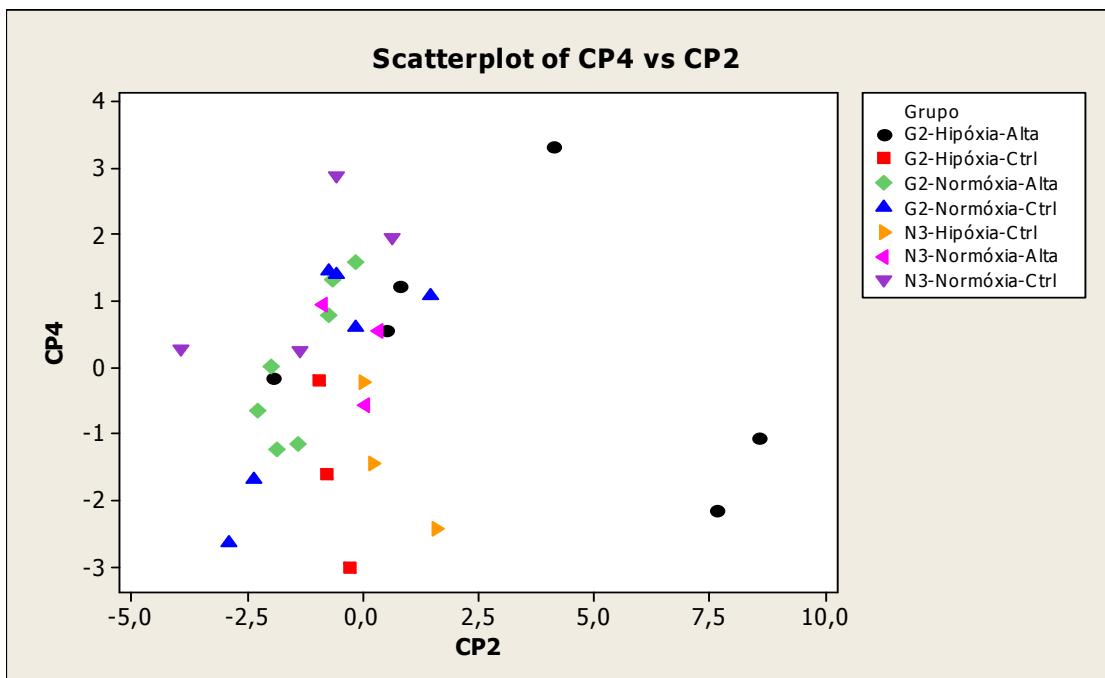


Gráfico 2.5: Dispersão: Componente 4 vs Componente 1





Apêndice 3: ANOVA - Análise de Variância

3.1)

Modelo de ANOVA adotado na análise dos dados de cada Metabólito:

$$Y_{ijkl} = \mu + \alpha_i + \beta_j + \varphi_k + (\alpha\beta)_{ij} + (\alpha\varphi)_{ik} + (\beta\varphi)_{jk} + \varepsilon_{ijkl} ;$$

- Com $i = 1, 2; j = 1, 2; k = 1, 2; l = 1, \dots, n$ (n é o número de réplicas em cada combinação)
- Com a restrição:

$$\sum_{i=1}^2 \alpha_i = \sum_{j=1}^2 \beta_j = \sum_{k=1}^2 \varphi_k = \sum_{i,j=1}^2 (\alpha\beta)_{ij} = \sum_{i,k=1}^2 (\alpha\varphi)_{ik} = \sum_{j,k=1}^2 (\beta\varphi)_{jk} = 0$$

- Variável Resposta:

Y_{ijkl} : valor para a l -ésima réplica da variável transformada, vinda da Linhagem i (G2, se $i = 1$; N3, se $i = 2$), submetida à condição de oxigênio j (Normoxia, se $j = 1$; Hipoxia, se $j = 2$) e ao nível de Glicose k (Alta, se $k = 1$; Controle, se $i = 2$)

- Os parâmetros mostrados podem ser interpretados da seguinte maneira:
 μ : média geral para a variável transformada, independentemente da condição de Linhagem, Oxigênio e Glicose à qual ela foi submetida.

α_i : incremento na média geral das respostas devido ao fato da célula ter vindo da Linhagem i .

β_j : incremento na média geral das respostas devido ao fato da célula ser submetida à condição de Oxigênio j .

φ_k : incremento na média geral das respostas devido ao fato da célula ser cultivada no nível de Glicose k .

$(\alpha\beta)_{ij}$: efeito de interação entre os fatores de Linhagem e Oxigênio.

$(\alpha\varphi)_{ik}$: efeito de interação entre os fatores de Linhagem e Glicose.

$(\beta\varphi)_{jk}$: efeito de interação entre os fatores de Oxigênio e Glicose.

ε_{ijkl} : erro aleatório da l -ésima observação.

- *Suposição: $\varepsilon_{ijkl} \sim \text{Normal}(0, \sigma^2)$, independentes*

3.2)

Na tabela, os campos e suas referências são:

- “Metabólitos” – variável analisada, de P1, P2, ... Poutros.
- “Ajuste” – ‘OK’ quando o ajuste foi feito diretamente por ANOVA; isto é, por mínimos quadrados ordinários ‘MQP’ se foi necessária a correção por Mínimos Quadrados Ponderados.
- “Interação de 1^a Ordem” – Se houver alguma interação significa que a presença comanda de 2 fatores altera a quantidade do metabólito presente na célula. Por exemplo, se a interação for entre Oxigênio e Glicose, quer dizer que há pelo menos uma diferença nas concentrações do metabólito entre os níveis Normóxia-Controle, Normóxia-Alta, Hipóxia-Controle e Hipóxia-Alta.
- “Linhagem” – ‘SIM’ quer dizer que existe diferença significante na concentração média do metabólito analisado quando comparamos G2 com N3.
- “Oxigênio” – ‘SIM’ quer dizer que existe diferença significante na concentração média do metabólito analisado quando comparamos Normóxia com Hipóxia.
- “Glicose” – ‘SIM’ quer dizer que existe diferença significante na concentração média do metabólito analisado quando comparamos Alta com Controle.

- **Grupo 1 – Distinção: NENHUMA**

Quantidade de Metabólitos: 8

Metabólitos: P2, P7, P15, P20, P22, P28, P30, Poutros

Tabela 3.33 : Efeito dos Fatores : NENHUM

ANOVA Dados Composicionais					
<i>Metabólitos</i>	<i>Ajuste?</i>	<i>Interação 1^a</i>	<i>Linhagem</i>	<i>Oxigênio</i>	<i>Glicose</i>
		<i>Ordem</i>			
P2	Ok	NÃO	NÃO	NÃO	NÃO
P7	OK	NÃO	NÃO	NÃO	NÃO
P15	Ok	NÃO	NÃO	NÃO	NÃO
P20	Ok	NÃO	NÃO	NÃO	NÃO
P22	Ok	NÃO	NÃO	NÃO	NÃO
P28	MQP	NÃO	NÃO	NÃO	NÃO
P30	MQP	NÃO	NÃO	NÃO	NÃO
Poutros	OK	NÃO	NÃO	NÃO	NÃO

- **Grupo 2 – Distinção: OXIGÊNIO**

Quantidade de Metabólitos: 4

Metabólitos: P6, P21, P24, P29

Tabela 3.34 : Efeito dos Fatores : Oxigênio

ANOVA Dados Composicionais					
<i>Metabólitos</i>	<i>Ajuste?</i>	<i>Interação 1^a Ordem</i>	<i>Linhagem</i>	<i>Oxigênio</i>	<i>Glicose</i>
P6	OK	NÃO	NÃO	SIM	NÃO
P21	OK	NÃO	NÃO	SIM	NÃO
P24	OK	NÃO	NÃO	SIM	NÃO
P29	OK	NÃO	NÃO	SIM	NÃO

- **Grupo 3: LINHAGEM**

Quantidade de Metabólitos: 1

Metabólitos: P3

- **Grupo 4: GLICOSE**

Quantidade de Metabólitos: 1

Metabólitos: P16

- **Grupo 5: LINHAGEM e OXIGÊNIO**

Quantidade de Metabólitos: 1

Metabólitos: P25

Tabela 3.35 : Efeito dos Fatores : Linhagem e Oxigênio

ANOVA Dados Composicionais					
<i>Metabólitos</i>	<i>Ajuste?</i>	<i>Interação 1^a Ordem</i>	<i>Linhagem</i>	<i>Oxigênio</i>	<i>Glicose</i>
P3	Ok	NÃO	SIM	NÃO	NÃO
P16	MQP	NÃO	NÃO	NÃO	SIM
P25	OK	NÃO	SIM	SIM	NÃO

- **Grupo 6 – Distinção: Interação OXIGÊNIO-GLICOSE**

Quantidade de Metabólitos: 7

Metabólitos: P4, P12, P13, P18, P26, P27, P31

Tabela 3.36 : Efeito de Interação entre os Fatores : Oxigênio e Glicose

ANOVA Dados Composicionais					
<i>Metabólitos</i>	<i>Ajuste?</i>	<i>Interação 1^a Ordem</i>	<i>Linhagem</i>	<i>Oxigênio</i>	<i>Glicose</i>
P4	MQP	Oxigênio/Glicose	NÃO	-	-
P12	OK	Oxigênio/Glicose	NÃO	-	-
P13	OK	Oxigênio/Glicose	NÃO	-	-
P18	OK	Oxigênio/Glicose	NÃO	-	-
P26	OK	Oxigênio/Glicose	NÃO	-	-
P27	Ok	Oxigênio/Glicose	NÃO	-	-
P31	OK	Oxigênio/Glicose	NÃO	-	-

- **Grupo 7 – Distinção: Interação LINHAGEM-GLICOSE**

Quantidade de Metabólitos: 2

Metabólitos: P8, P11

Tabela 3.37 : Efeito de Interação entre os Fatores : Linhagem e Glicose

ANOVA Dados Composicionais					
<i>Metabólitos</i>	<i>Ajuste?</i>	<i>Interação 1^a Ordem</i>	<i>Linhagem</i>	<i>Oxigênio</i>	<i>Glicose</i>
P8	MQP	Linhagem /Glicose	-	NÃO	-
P11	Ok	Linhagem/Glicose	-	SIM	-

- **Grupo 8 – Distinção: Interação OXIGÊNIO-GLICOSE e LINHAGEM-GLICOSE**

Quantidade de Metabólitos: 3

Metabólitos: P5, P9, P17

Tabela 3.38 : Efeito de Interação entre os Fatores : Oxigenio e Glicose, Linhagem e Glicose

ANOVA Dados Composicionais					
Metabólitos	Ajuste?	Interação 1 ^a Ordem	Linhagem	Oxigênio	Glicose
P5	OK	Oxigênio/Glicose - Linhagem /Glicose	-	-	-
P9	MQP	Oxigênio/Glicose - Linhagem /Glicose	-	-	-
P17	OK	Oxigênio/Glicose - Linhagem /Glicose	-	-	-

- **Grupo 9 – Distinção: Interação LINHAGEM-OXIGÊNIO e LINHAGEM-GLICOSE**

Quantidade de Metabólitos: 2

Metabólitos: P10, P23

Tabela 3.39 : Efeito de Interação entre os Fatores : Linhagem e Oxigênio, Linhagem e Glicose

ANOVA Dados Composicionais					
Metabólitos	Ajuste?	Interação 1 ^a Ordem	Linhagem	Oxigênio	Glicose
P10	MQP	Linhagem /Glicose - Linhagem /Oxigênio	-	-	-
P23	OK	Linhagem /Glicose - Linhagem Oxigênio	-	-	-

- **Grupo 10 – Distinção: Interações - TODAS**

Quantidade de Metabólitos: 3

Metabólitos: P1, P14, P19

Tabela 3.40 : Efeito de todas as Interações

ANOVA Dados Composicionais					
Metabólitos	Ajuste?	Interação 1 ^a Ordem	Linhagem	Oxigênio	Glicose
P1	MQP	TODAS	-	-	-
P14	OK	TODAS	-	-	-
P19	MQP	TODAS	-	-	-

3.3) Direção dos Efeitos

Tabela 3.3: Direção dos Efeitos Principais

Tabela de Comparações			
METABÓLITOS	COMBINAÇÃO	SINAL(<i>p</i> -value)	COMBINAÇÃO
<i>Efeito de Oxigênio</i>			
P6	Média Normoxia	>(0.006)	Média Hipoxia
P21	Média Normoxia	<(0.0402)	Média Hipoxia
P24	Média Normoxia	<(0.0196)	Média Hipoxia
P29	Média Normoxia	>(0.0172)	Média Hipoxia
<i>Efeito de Linhagem</i>			
P3	Média G2	<(0.0049)	Média N3
<i>Efeito de Glicose</i>			
P16	Média Controle	<(0.0486)	Média Alta
<i>Efeito de Linhagem e Oxigênio</i>			
P25	Média G2	<(0.0214)	Média N3
P25	Média Normoxia	<(0.0222)	Média Hipoxia

Tabela 3.4: Direção dos Efeitos de Interação

Tabela de Comparações Múltiplas			
METABÓLITOS	COMBINAÇÃO	SINAL	
<i>Interação Oxigênio/Glicose</i>			
P4	Média Normoxia-Controle	>	Média Hipoxia-Controle
P4	Média Normoxia-Controle	>	Média Normoxia-Alta
P4	Média Hipoxia-Alta	>	Média Hipoxia-Controle
P4	Média Normoxia-Alta	<	Média Hipoxia-Alta
-	-	-	-
P12	Média Normoxia-Controle	=	Média Hipoxia-Controle
P12	Média Normoxia-Controle	<	Média Normoxia-Alta
P12	Média Hipoxia-Alta	<	Média Hipoxia-Controle
P12	Média Normoxia-Alta	>	Média Hipoxia-Alta
-	-	-	-
P13	Média Normoxia-Controle		Média Hipoxia-Controle
P13	Média Normoxia-Controle		Média Normoxia-Alta
P13	Média Hipoxia-Alta		Média Hipoxia-Controle
P13	Média Normoxia-Alta		Média Hipoxia-Alta
-	-	-	-
P18	Média Normoxia-Controle	<	Média Hipoxia-Controle
P18	Média Normoxia-Controle	=	Média Normoxia-Alta
P18	Média Hipoxia-Alta	<	Média Hipoxia-Controle

P18	Média Normóxia-Alta	>	Média Hipóxia-Alta
-	-	-	-
P26	Média Normóxia-Controle	<	Média Hipóxia-Controle
P26	Média Normóxia-Controle	=	Média Normóxia-Alta
P26	Média Hipóxia-Alta	<	Média Hipóxia-Controle
P26	Média Normóxia-Alta	>	Média Hipóxia-Alta
-	-	-	-
P27	Média Normóxia-Controle	<	Média Hipóxia-Controle
P27	Média Normóxia-Controle	<	Média Normóxia-Alta
P27	Média Hipóxia-Alta		Média Hipóxia-Controle
P27	Média Normóxia-Alta	>	Média Hipóxia-Alta
-	-	-	-
P31	Média Normóxia-Controle	>	Média Hipóxia-Controle
P31	Média Normóxia-Controle	=	Média Normóxia-Alta
P31	Média Hipóxia-Alta	<	Média Hipóxia-Controle
P31	Média Normóxia-Alta	>	Média Hipóxia-Alta

Tabela de Comparações Múltiplas			
METABÓLITOS	COMBINAÇÃO	SINAL	
<i>Interação Linhagem-Glicose</i>			
P8	Média G2-Controle	>	Média N3-Controle
P8	Média G2-Alta	=	Média N3-Alta
P8	Média N3-Controle	<	Média N3-Alta
P8	Média G2-Alta	<	Média G2-Controle
P11	Média G2-Controle	>	Média N3-Controle
P11	Média G2-Alta	>	Média N3-Alta
P11	Média N3-Controle	<	Média N3-Alta
P11	Média G2-Alta	=	Média G2-Controle

Observação:

Por motivo de simplicidade nas interpretações não se apresentou os níveis descriptivos da comparação de Tukey para os efeitos de interações.

3.4) Diagnóstico de Ajuste

- ANOVA para o Metabólito P1 – Mínimos Quadrados Generalizados

Fonte	DF	Seq SS	Adj SS	Adj MS	F	P
Linhagem	1	34,33	1,81	1,81	2,05	0,165
Oxigênio	1	63,34	1,89	1,89	2,14	0,156
Glicose	1	53,65	3,41	3,41	3,85	0,061
Linhagem*Oxigênio	1	201,49	302,73	302,73	341,96	0
Linhagem*Glicose	1	814,00	315,02	315,02	355,84	0
Oxigênio*Glicose	1	587,42	587,42	587,42	663,54	0
Erro	25	22,13	22,13	0,89		
Total	31	1776,36				

$$S = 0,940892 \quad R-Sq = 98,75\% \quad R-Sq(adj) = 98,46\%$$

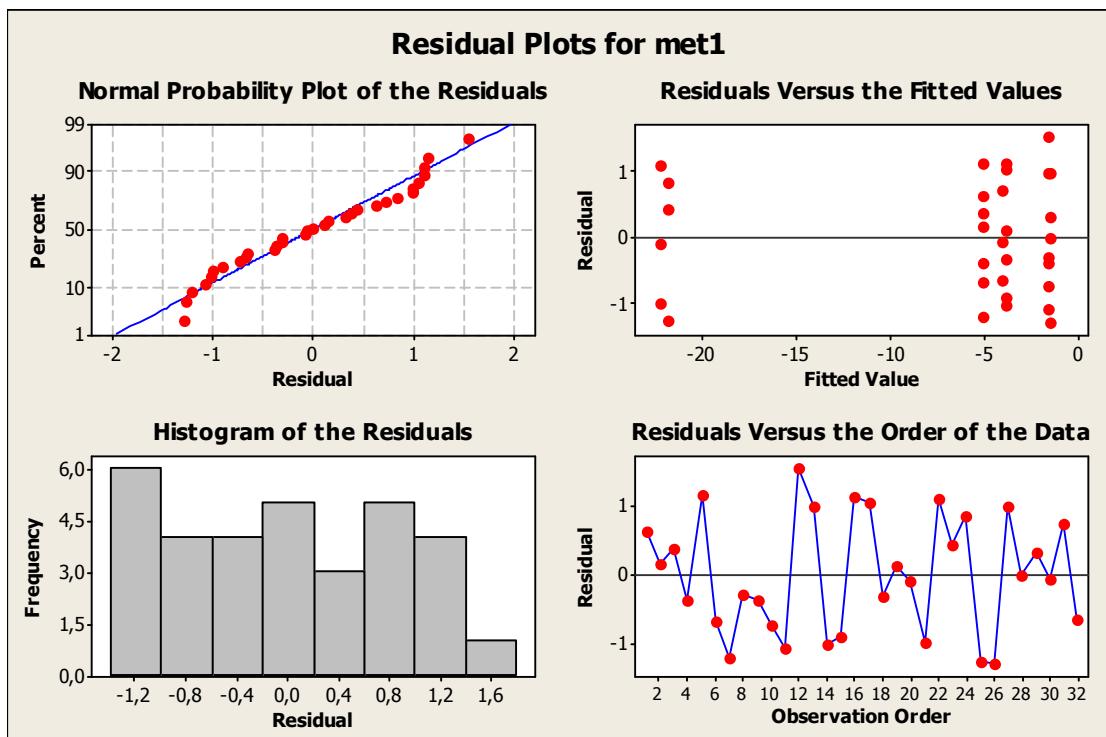


Gráfico 3.1 : Gráficos de Diagnóstico do Modelo

- ANOVA para o Metabólito P2

Fonte	DF	Seq SS	Adj SS	Adj MS	F	P
Linhagem	1	0,0559	0,2085	0,2085	1,94	0,176
Oxigênio	1	0,4	0,0975	0,0975	0,91	0,350
Glicose	1	0,1706	0,0024	0,0024	0,02	0,882
Linhagem*Oxigênio	1	0,0876	0,2997	0,2997	2,79	0,107
Linhagem*Glicose	1	0,0138	0,0145	0,0145	0,13	0,717
Oxigênio*Glicose	1	0,3984	0,3984	0,3984	3,71	0,065
Erro	25	2,6828	2,6828	0,1073		
Total	31	3,4095				

$$S = 0,327584 \quad R-Sq = 21,32\% \quad R-Sq(adj) = 2,43\%$$

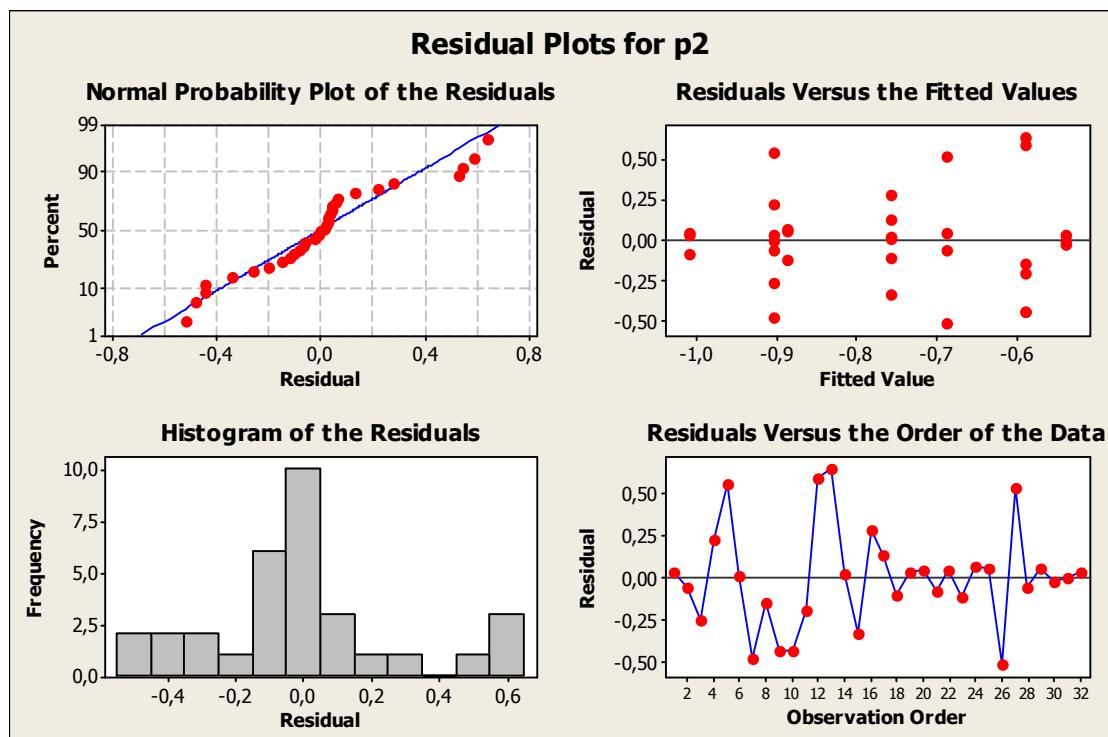


Gráfico 3.2 : Gráficos de Diagnóstico do Modelo

- ANOVA para o Metabólito P3

Fonte	DF	Seq SS	Adj SS	Adj MS	F	P
Linhagem	1	0,98066	0,93804	0,93804	9,62	0,005
Oxigênio	1	0,05404	0,14921	0,14921	1,53	0,228
Glicose	1	0,07394	0,17047	0,17047	1,75	0,198
Linhagem*Oxigênio	1	0,01555	0,05103	0,05103	0,52	0,476
Linhagem*Glicose	1	0,16154	0,15375	0,15375	1,58	0,221
Oxigênio*Glicose	1	0,00238	0,00238	0,00238	0,02	0,877
Erro	25	2,43717	2,43717	0,09749		
Total	31	3,72529				

$$S = 0,312229 \quad R-Sq = 34,58\% \quad R-Sq(adj) = 18,88\%$$

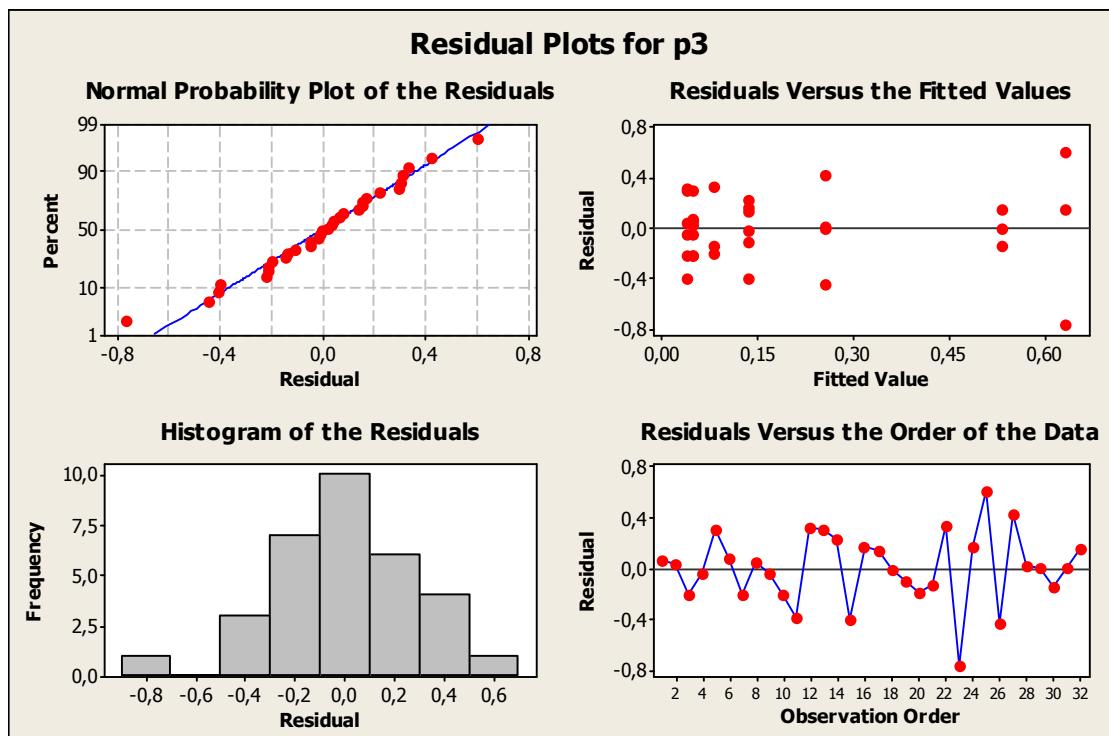


Gráfico 3.4 : Gráficos de Diagnóstico do Modelo

- ANOVA para o Metabólito P4 – Mínimos Quadrados Generalizados

Fonte	DF	Seq SS	Adj SS	Adj MS	F	P
Linhagem	1	0,132	0,576	0,576	0,57	0,459
Oxigênio	1	0,105	0,892	0,892	0,88	0,358
Glicose	1	0,400	0,180	0,180	0,18	0,678
Linhagem*Oxigênio	1	0,583	3,175	3,175	3,12	0,090
Linhagem*Glicose	1	5,679	1,052	1,052	1,03	0,319
Oxigênio*Glicose	1	10,754	10,754	10,754	10,56	0,003
Erro	25	25,468	25,468	1,019		
Total	31	43,120				

$$S = 1,00931 \quad R-Sq = 40,94\% \quad R-Sq(adj) = 26,76\%$$

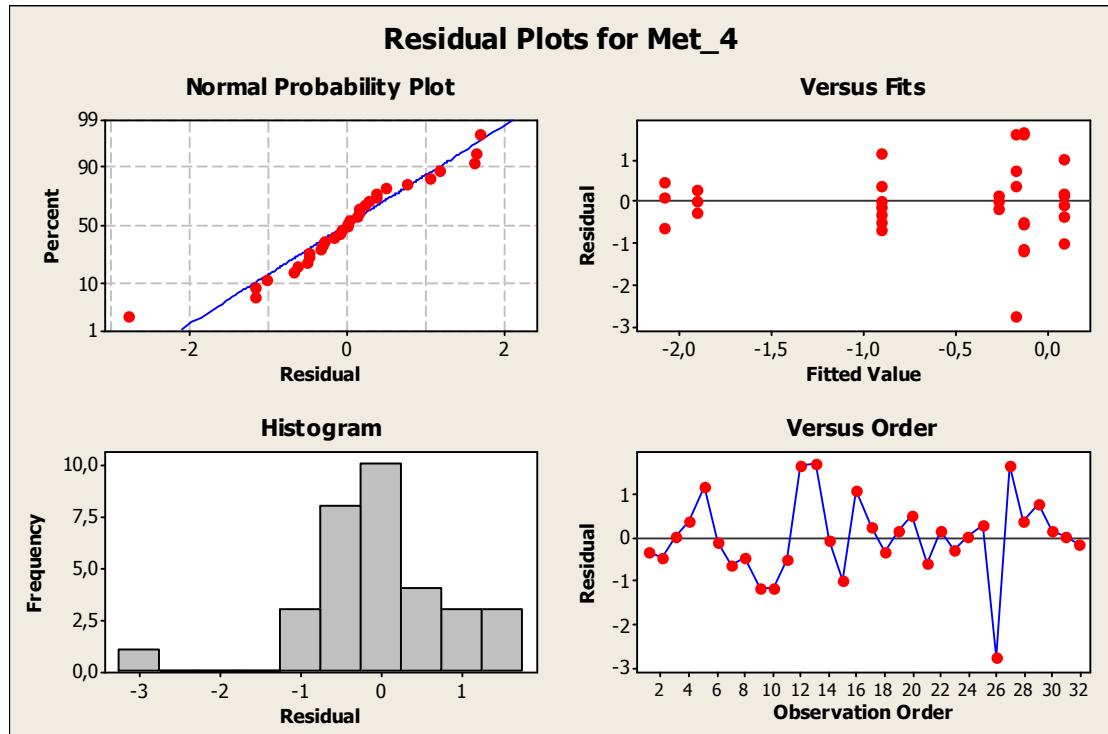


Gráfico 3.4 : Gráficos de Diagnóstico do Modelo

- ANOVA para o Metabólito P5

Fonte	DF	Seq SS	Adj SS	Adj MS	F	P
Linhagem	1	0,01056	0,00358	0,00358	0,12	0,731
Oxigênio	1	0,30048	0,28751	0,28751	9,72	0,005
Glicose	1	0,19034	0,10967	0,10967	3,71	0,066
Linhagem*Oxigênio	1	0,05195	0,00297	0,00297	0,10	0,754
Linhagem*Glicose	1	0,37480	0,14355	0,14355	4,86	0,037
Oxigênio*Glicose	1	0,27613	0,27613	0,27613	9,34	0,005
Erro	25	0,73916	0,73916	0,02957		
Total	31	1,94341				

$$S = 0,171948 \quad R-Sq = 61,97\% \quad R-Sq(adj) = 52,84\%$$

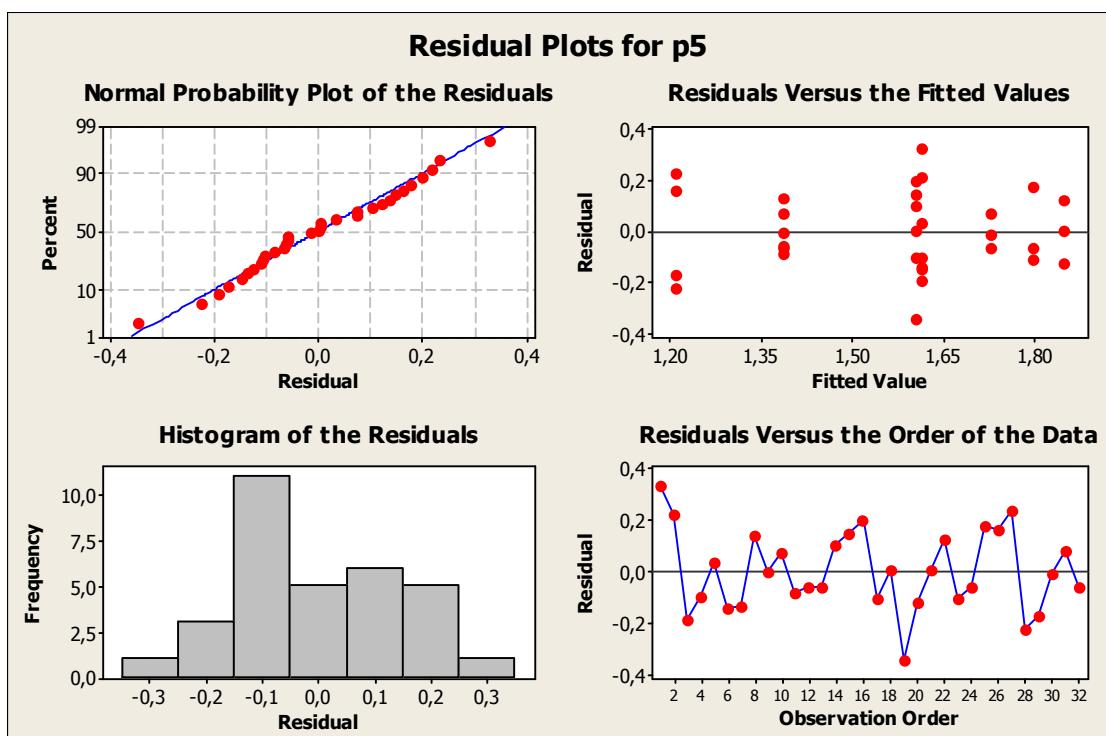


Gráfico 3.5 : Gráficos de Diagnóstico do Modelo

- ANOVA para o Metabólito P6

Fonte	DF	Seq SS	Adj SS	Adj MS	F	P
Linhagem	1	0,00622	0,01279	0,01279	0,35	0,558
Oxigênio	1	0,52494	0,53318	0,53318	14,70	0,001
Glicose	1	0,09690	0,41	0,41	0,01	0,916
Linhagem*Oxigênio	1	0,01096	0,04679	0,04679	1,29	0,267
Linhagem*Glicose	1	0,07164	0,08650	0,08650	2,38	0,135
Oxigênio*Glicose	1	0,01515	0,01515	0,01515	0,42	0,524
Erro	25	0,90703	0,90703	0,03628		
Total	31	1,63285				

$$S = 0,190477 \quad R-Sq = 44,45\% \quad R-Sq(adj) = 31,12\%$$

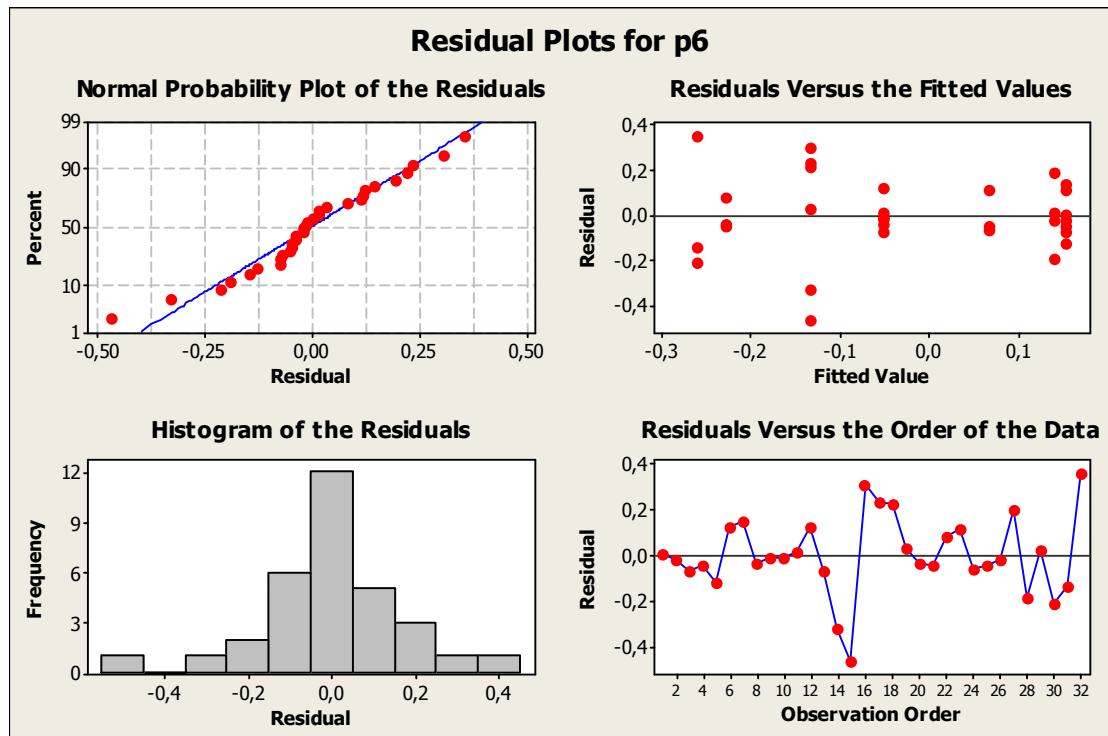


Gráfico 3.6 : Gráficos de Diagnóstico do Modelo

- ANOVA para o Metabólito P7

Fonte	DF	Seq SS	Adj SS	Adj MS	F	P
Linhagem	1	0,1657	0,0478	0,0478	0,34	0,563
Oxigênio	1	0,1231	0,0132	0,0132	0,09	0,761
Glicose	1	0,0081	0,0276	0,0276	0,20	0,660
Linhagem*Oxigênio	1	0,0556	0,0200	0,0200	0,14	0,708
Linhagem*Glicose	1	0,3578	0,1789	0,1789	1,28	0,268
Oxigênio*Glicose	1	0,1357	0,1357	0,1357	0,97	0,333
Erro	25	3,4833	3,4833	0,1393		
Total	31	4,3293				

$$S = 0,373271 \quad R-Sq = 19,54\% \quad R-Sq(adj) = 0,23\%$$

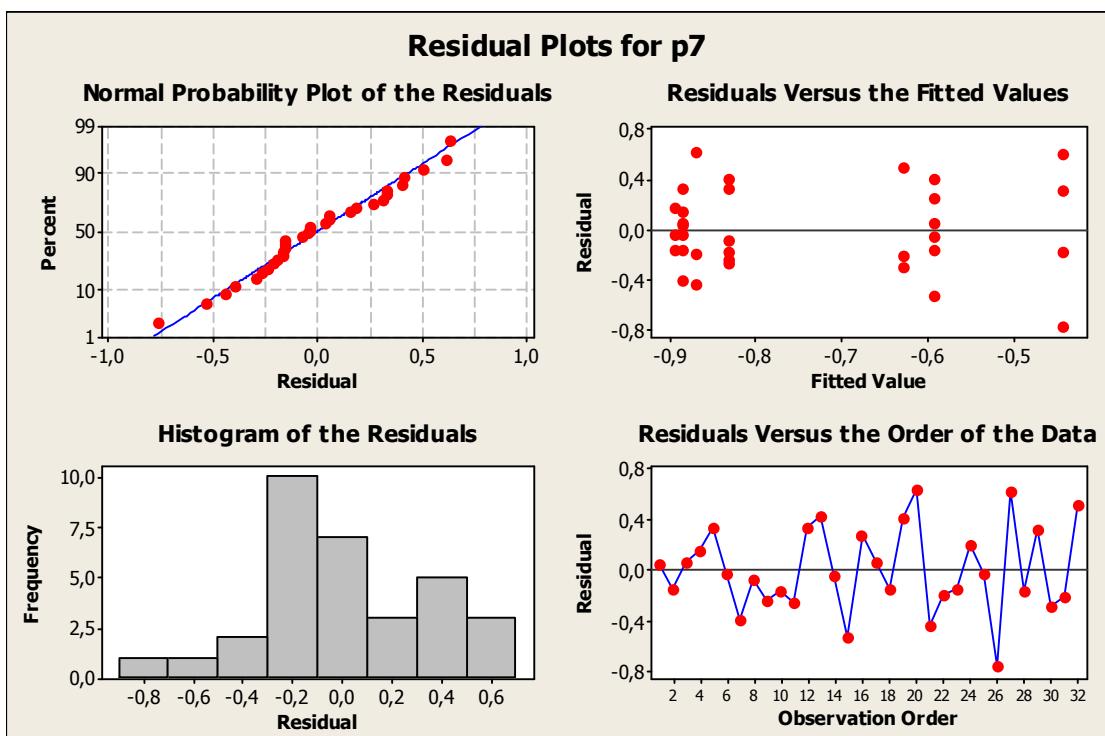


Gráfico 3.7 : Gráficos de Diagnóstico do Modelo

- ANOVA para o Metabólito P8 – Mínimos Quadrados Generalizados

Fonte	DF	Seq SS	Adj SS	Adj MS	F	P
Linhagem	1	2,852	2,760	2,760	2,73	0,111
Oxigênio	1	0,018	0,171	0,171	0,17	0,684
Glicose	1	0,171	1,012	1,012	1,00	0,327
Linhagem*Oxigênio	1	0,098	0	0	0,00	0,987
Linhagem*Glicose	1	7,609	5,244	5,244	5,18	0,032
Oxigênio*Glicose	1	0,588	0,588	0,588	0,58	0,453
Erro	25	25,307	25,307	1,012		
Total	31	36,643				

$S = 1,00613$ $R-Sq = 30,94\%$ $R-Sq(adj) = 14,36\%$

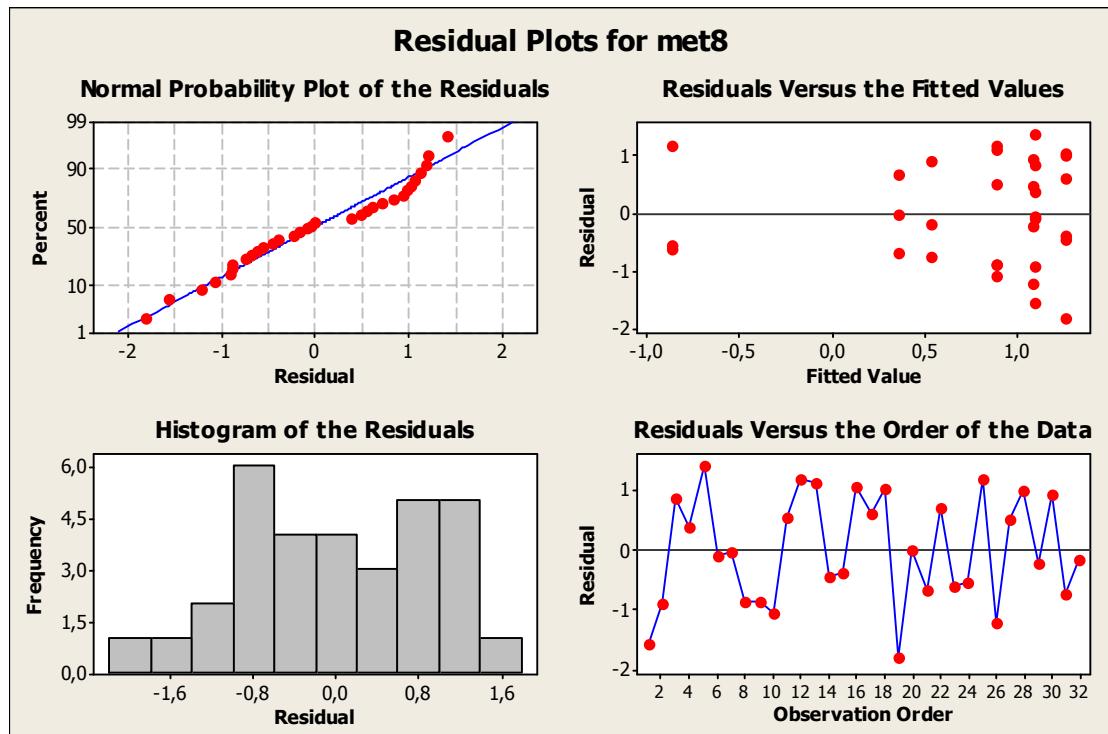


Gráfico 3.8 : Gráficos de Diagnóstico do Modelo

- ANOVA para o Metabólito P9 – Mínimos Quadrados Generalizados

Fonte	DF	Seq SS	Adj SS	Adj MS	F	P
Linhagem	1	0,0145	0,0840	0,0840	0,10	0,759
Oxigênio	1	0,8259	0,3649	0,3649	0,42	0,524
Glicose	1	0,5678	0,2206	0,2206	0,25	0,620
Linhagem*Oxigênio	1	0,5250	1,1519	1,1519	1,32	0,262
Linhagem*Glicose	1	9,3894	4,0295	4,0295	4,61	0,042
Oxigênio*Glicose	1	5,4066	5,4066	5,4066	6,18	0,020
Erro	25	21,8646	21,8646	0,8746		
Total	31	38,5939				

$$S = 0,935192 \quad R-Sq = 43,35\% \quad R-Sq(adj) = 29,75\%$$

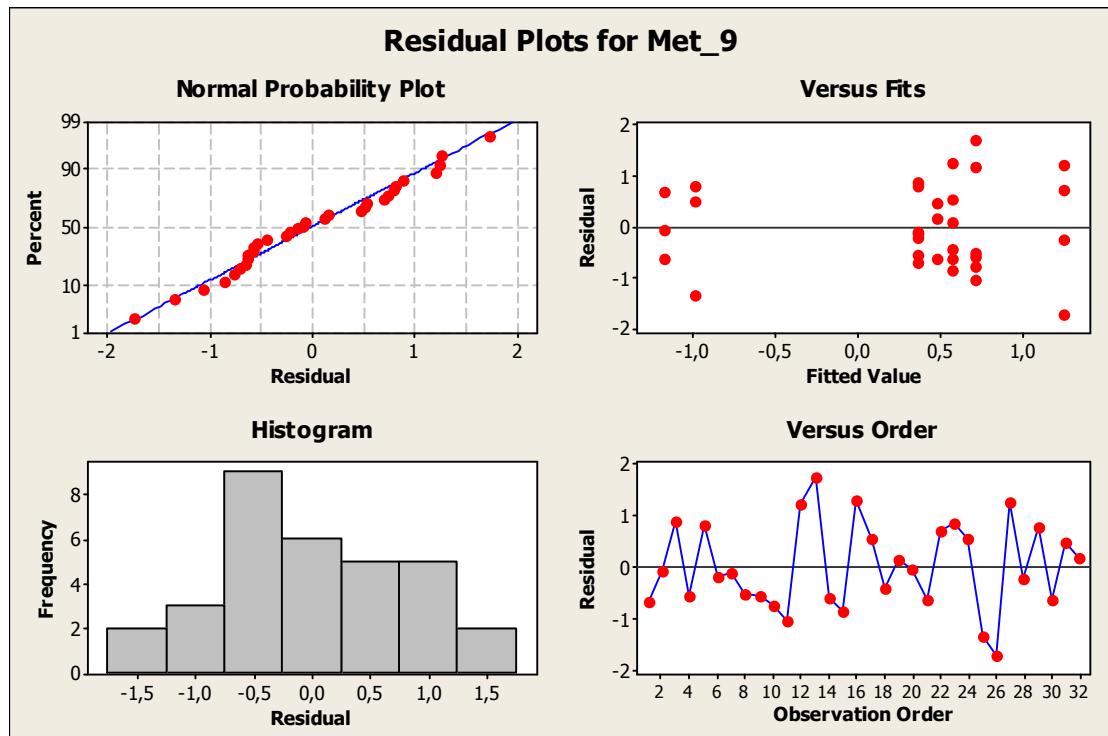


Gráfico 3.9 : Gráficos de Diagnóstico do Modelo

- ANOVA para o Metabólito P10 – Mínimos Quadrados Generalizados

Fonte	DF	Seq SS	Adj SS	Adj MS	F	P
Linhagem	1	1,3027	2,8323	2,8323	3,49	0,073
Oxigênio	1	0,5044	1,4820	1,4820	1,83	0,189
Glicose	1	1,6816	0,0076	0,0076	0,01	0,923
Linhagem*Oxigênio	1	0,6530	3,0348	3,0348	3,74	0,064
Linhagem*Glicose	1	5,0383	0,9441	0,9441	1,16	0,291
Oxigênio*Glicose	1	9,4457	9,4457	9,4457	11,65	0,002
Erro	25	20,2767	20,2767	0,8111		
Total	31	38,9025				

$$S = 0,900593 \quad R-Sq = 47,88\% \quad R-Sq(adj) = 35,37\%$$

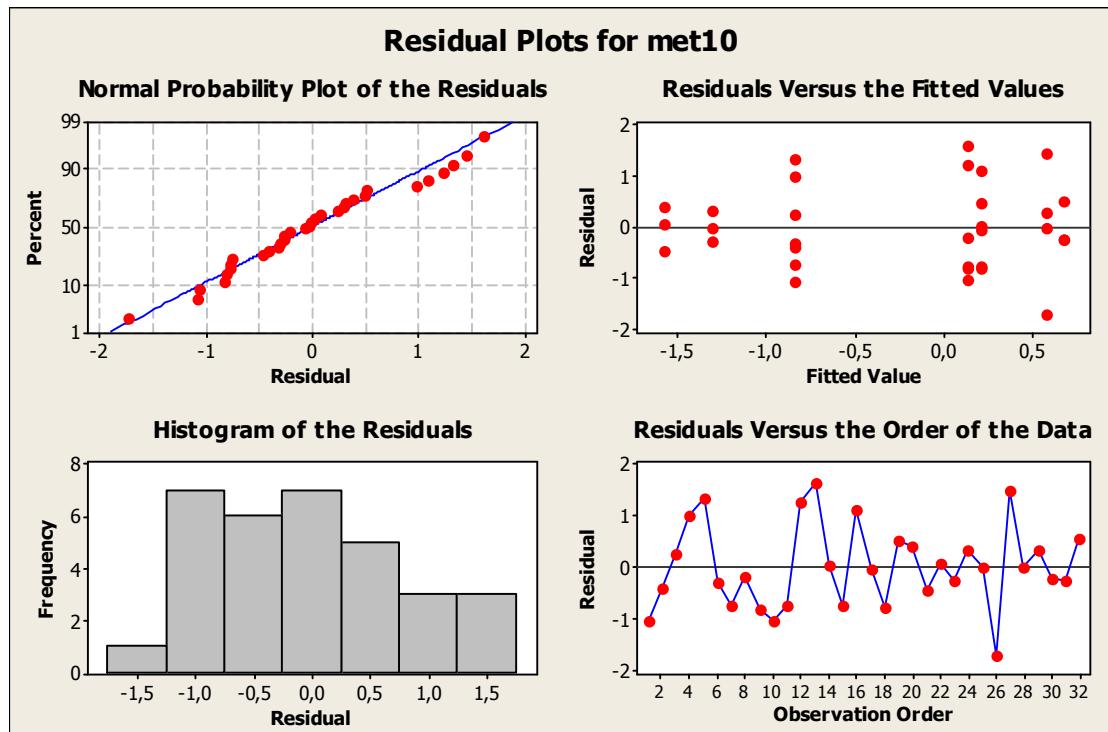


Gráfico 3.10 : Gráficos de Diagnóstico do Modelo

- ANOVA para o Metabólito P11

Fonte	DF	Seq SS	Adj SS	Adj MS	F	P
Linhagem	1	0,03858	0,09238	0,09238	2,16	0,154
Oxigênio	1	0,60044	0,78296	0,78296	18,28	0
Glicose	1	0,00342	0,18320	0,18320	4,28	0,049
Linhagem*Oxigênio	1	0,00543	0,19610	0,19610	4,58	0,042
Linhagem*Glicose	1	0,00301	0,04814	0,04814	1,12	0,299
Oxigênio*Glicose	1	0,55214	0,55214	0,55214	12,89	0,001
Erro	25	1,07096	1,07096	0,04284		
Total	31	2,27399				

$$S = 0,206974 \quad R-Sq = 52,90\% \quad R-Sq(adj) = 41,60\%$$

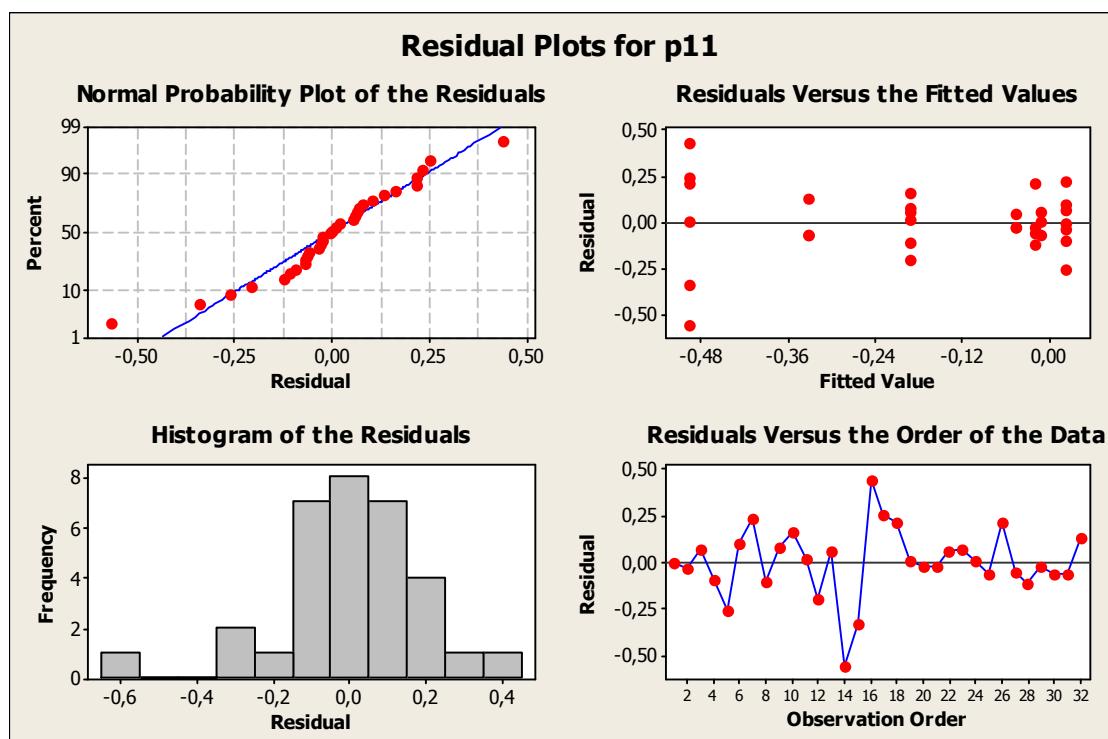


Gráfico 3.11 : Gráficos de Diagnóstico do Modelo

- ANOVA para o Metabólito P12

Fonte	DF	Seq SS	Adj SS	Adj MS	F	P
Linhagem	1	0,02	0,09215	0,09215	1,81	0,191
Oxigênio	1	0,26394	0,34296	0,34296	6,72	0,016
Glicose	1	0,51	0,06764	0,06764	1,33	0,261
Linhagem*Oxigênio	1	0,00985	0,11591	0,11591	2,27	0,144
Linhagem*Glicose	1	0,03534	0,00115	0,00115	0,02	0,882
Oxigênio*Glicose	1	0,32945	0,32945	0,32945	6,46	0,018
Erro	25	1,27588	1,27588	0,05104		
Total	31	1,91500				

$$S = 0,225910 \quad R-Sq = 33,37\% \quad R-Sq(adj) = 17,38\%$$

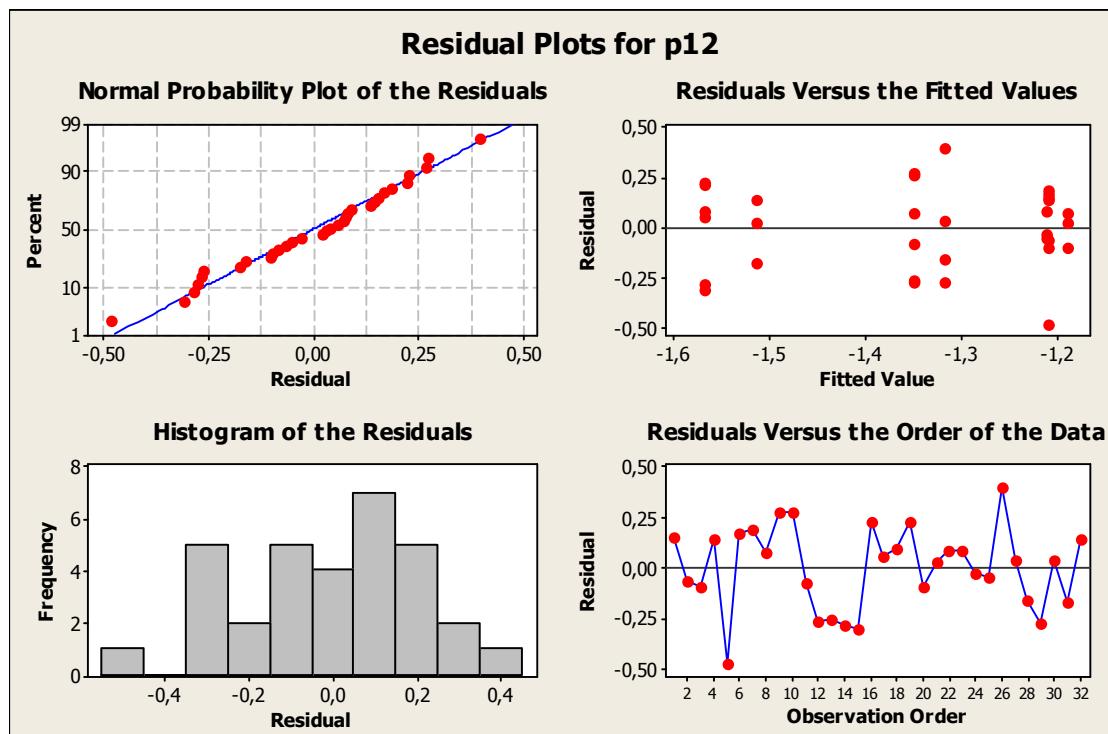


Gráfico 3.12 : Gráficos de Diagnóstico do Modelo

- ANOVA para o Metabólito P13

Fonte	DF	Seq SS	Adj SS	Adj MS	F	P
Linhagem	1	0,6	0,3488	0,3488	2,27	0,144
Oxigênio	1	0,7260	1,0018	1,0018	6,53	0,017
Glicose	1	0,1981	0,6735	0,6735	4,39	0,046
Linhagem*Oxigênio	1	0,0051	0,2030	0,2030	1,32	0,261
Linhagem*Glicose	1	0,0038	0,1549	0,1549	1,01	0,325
Oxigênio*Glicose	1	0,8537	0,8537	0,8537	5,57	0,026
Erro	25	3,8344	3,8344	0,1534		
Total	31	5,6218				

$$S = 0,391633 \quad R-Sq = 31,79\% \quad R-Sq(adj) = 15,42\%$$

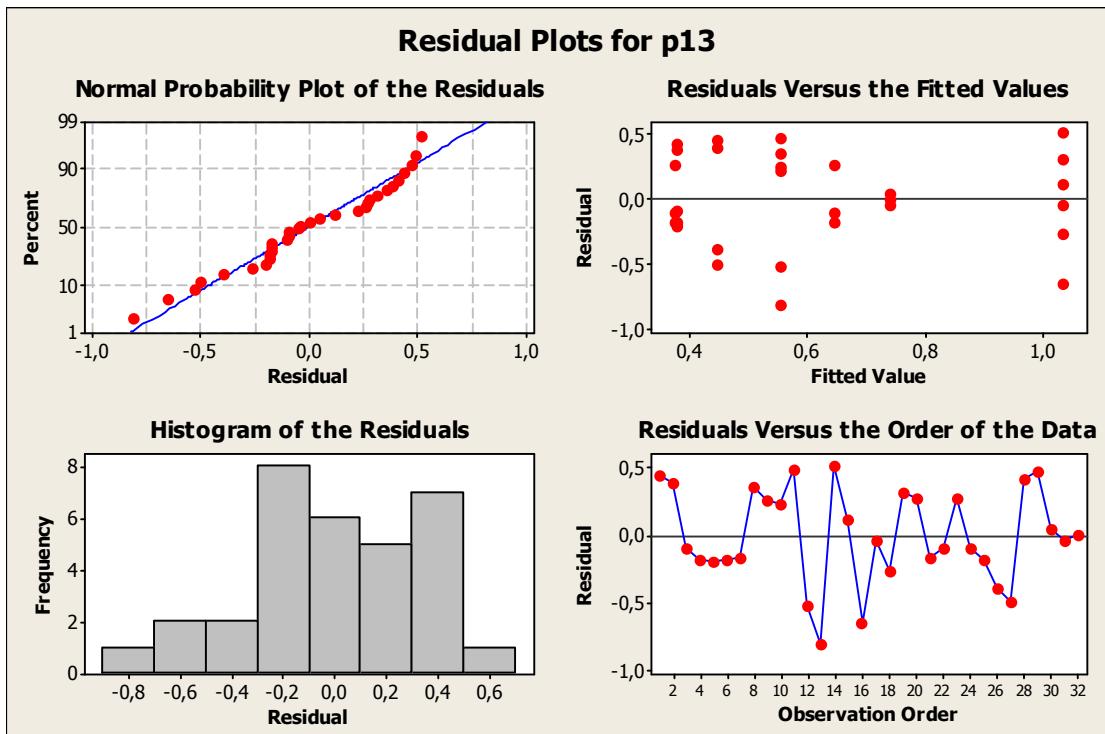


Gráfico 3.13 : Gráficos de Diagnóstico do Modelo

- ANOVA para o Metabólito P14

Fonte	DF	Seq SS	Adj SS	Adj MS	F	P
Linhagem	1	0,3834	0,3621	0,3621	2,81	0,106
Oxigênio	1	0,4174	0,0101	0,0101	0,08	0,781
Glicose	1	0,8193	0,0708	0,0708	0,55	0,465
Linhagem*Oxigênio	1	1,3528	0,6810	0,6810	5,29	0,030
Linhagem*Glicose	1	2,2936	1,1609	1,1609	9,02	0,006
Oxigênio*Glicose	1	0,8368	0,8368	0,8368	6,50	0,017
Erro	25	3,2165	3,2165	0,1287		
Total	31	9,3199				

$$S = 0,358692 \quad R-Sq = 65,49\% \quad R-Sq(adj) = 57,20\%$$

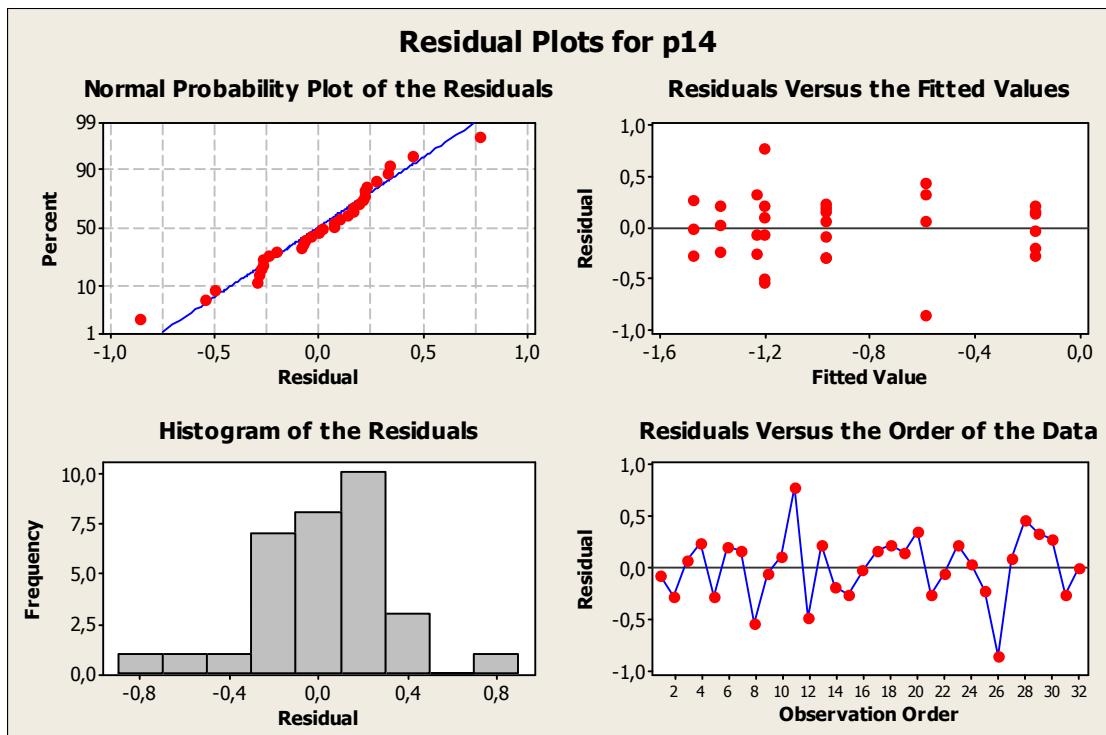


Gráfico 3.14 : Gráficos de Diagnóstico do Modelo

- ANOVA para o Metabólito P15

Fonte	DF	Seq SS	Adj SS	Adj MS	F	P
Linhagem	1	0,0706	0,1020	0,1020	0,28	0,600
Oxigênio	1	2,5617	0,7761	0,7761	2,15	0,155
Glicose	1	0,4134	0,0360	0,0360	0,10	0,755
Linhagem*Oxigênio	1	0,5947	0,6233	0,6233	1,73	0,201
Linhagem*Glicose	1	0,0552	0,0044	0,0044	0,01	0,913
Oxigênio*Glicose	1	0,1753	0,1753	0,1753	0,49	0,492
Erro	25	9,0298	9,0298	0,3612		
Total	31	12,9006				

$S = 0,600993$ $R-Sq = 30,00\%$ $R-Sq(\text{adj}) = 13,21\%$

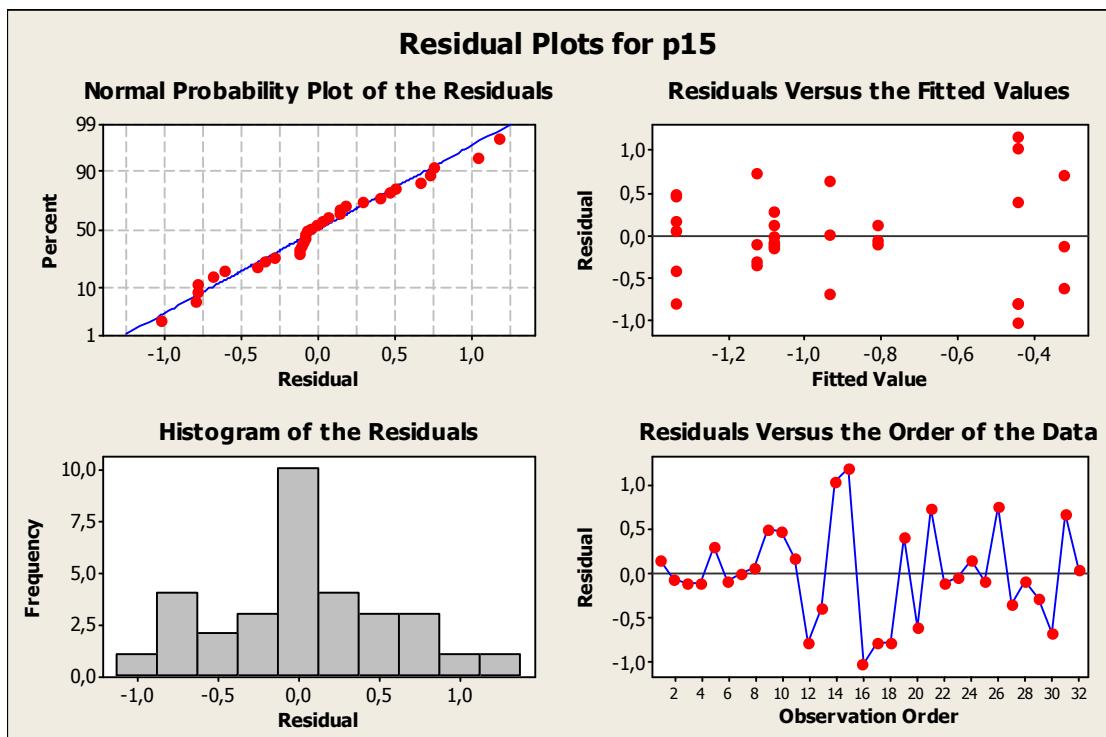


Gráfico 3.15 : Gráficos de Diagnóstico do Modelo

- ANOVA para o Metabólito P16 – Mínimos Quadrados Generalizados

Fonte	DF	Seq SS	Adj SS	Adj MS	F	P
Linhagem	1	45,626	40,109	40,109	44,46	0
Oxigênio	1	24,125	36,936	36,936	40,94	0
Glicose	1	1,493	0,590	0,590	0,65	0,426
Linhagem*Oxigênio	1	11,213	5,721	5,721	6,34	0,019
Linhagem*Glicose	1	53,810	28,996	28,996	32,14	0
Oxigênio*Glicose	1	15,804	15,804	15,804	17,52	0
Erro	25	22,555	22,555	0,902		
Total	31	174,627				

$S = 0,949832$ $R-Sq = 87,08\%$ $R-Sq(\text{adj}) = 83,98\%$

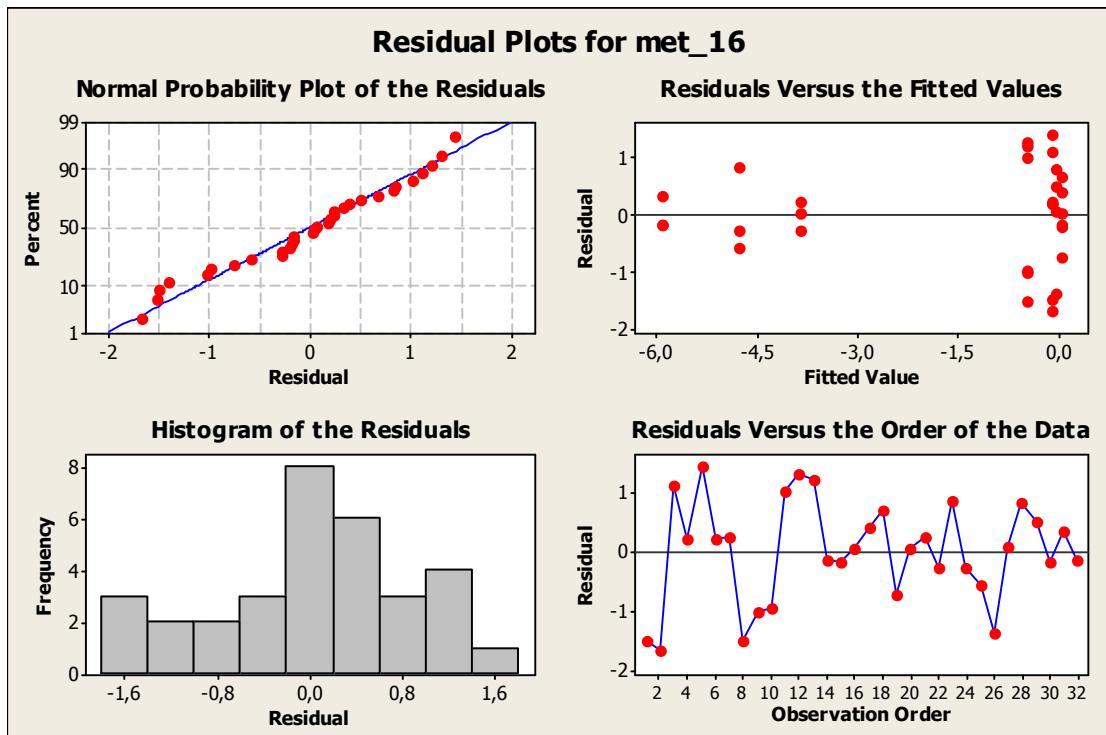


Gráfico 3.16 : Gráficos de Diagnóstico do Modelo

- ANOVA para o Metabólito P17

Fonte	DF	Seq SS	Adj SS	Adj MS	F	P
Linhagem	1	0,01051	0,04338	0,04338	0,61	0,442
Oxigênio	1	0,11457	0,13700	0,13700	1,92	0,178
Glicose	1	0,56706	0,23343	0,23343	3,28	0,082
Linhagem*Oxigênio	1	0,04031	0,03	0,03	0,00	0,985
Linhagem*Glicose	1	0,96931	0,40637	0,40637	5,71	0,025
Oxigênio*Glicose	1	0,58935	0,58935	0,58935	8,28	0,008
Erro	25	1,78035	1,78035	0,07121		
Total	31	4,07147				

$$S = 0,266859 \quad R-Sq = 56,27\% \quad R-Sq(adj) = 45,78\%$$

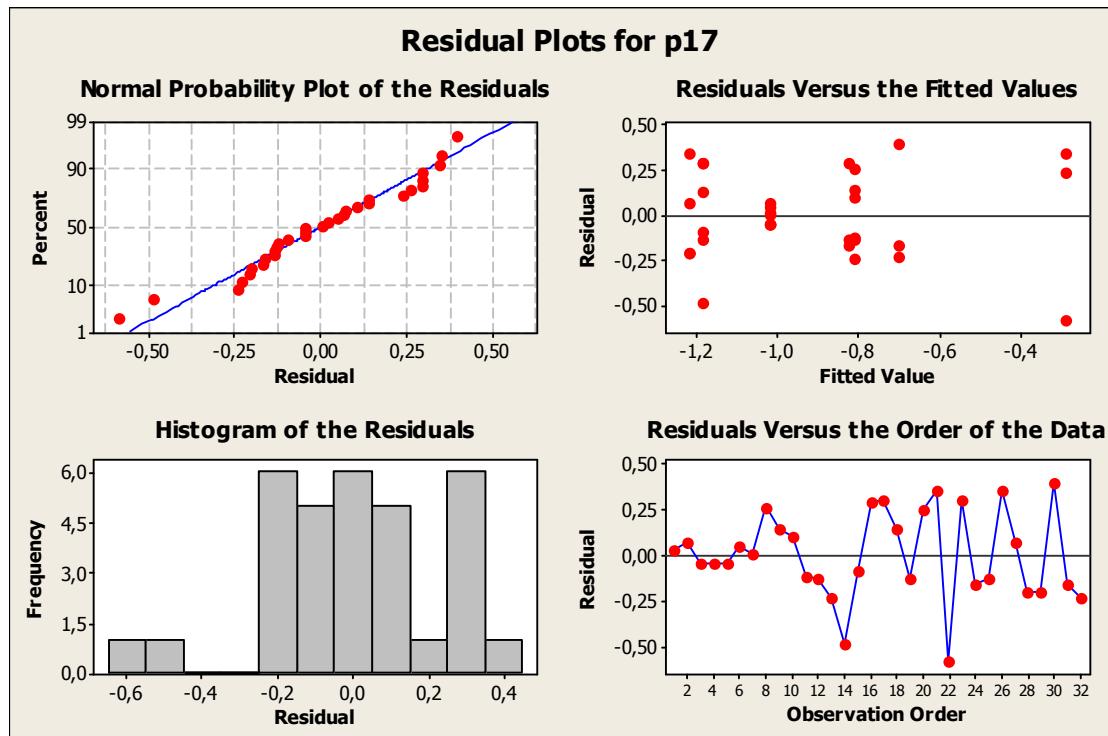


Gráfico 3.17 : Gráficos de Diagnóstico do Modelo

ANOVA para o Metabólito P18

Fonte	DF	Seq SS	Adj SS	Adj MS	F	P
Linhagem	1	0,0326	0,0	0,0 0,00	0,993	
Oxigênio	1	0,4037	0,0588	0,0588	0,26	0,617
Glicose	1	0,4495	0,1626	0,1626	0,71	0,407
Linhagem*Oxigênio	1	0,4009	0,0531	0,0531	0,23	0,634
Linhagem*Glicose	1	1,6792	0,6358	0,6358	2,78	0,108
Oxigênio*Glicose	1	1,2656	1,2656	1,2656	5,53	0,027
Erro	25	5,7197	5,7197	0,2288		
Total	31	9,9511				

$$S = 0,478318 \quad R-Sq = 42,52\% \quad R-Sq(adj) = 28,73\%$$

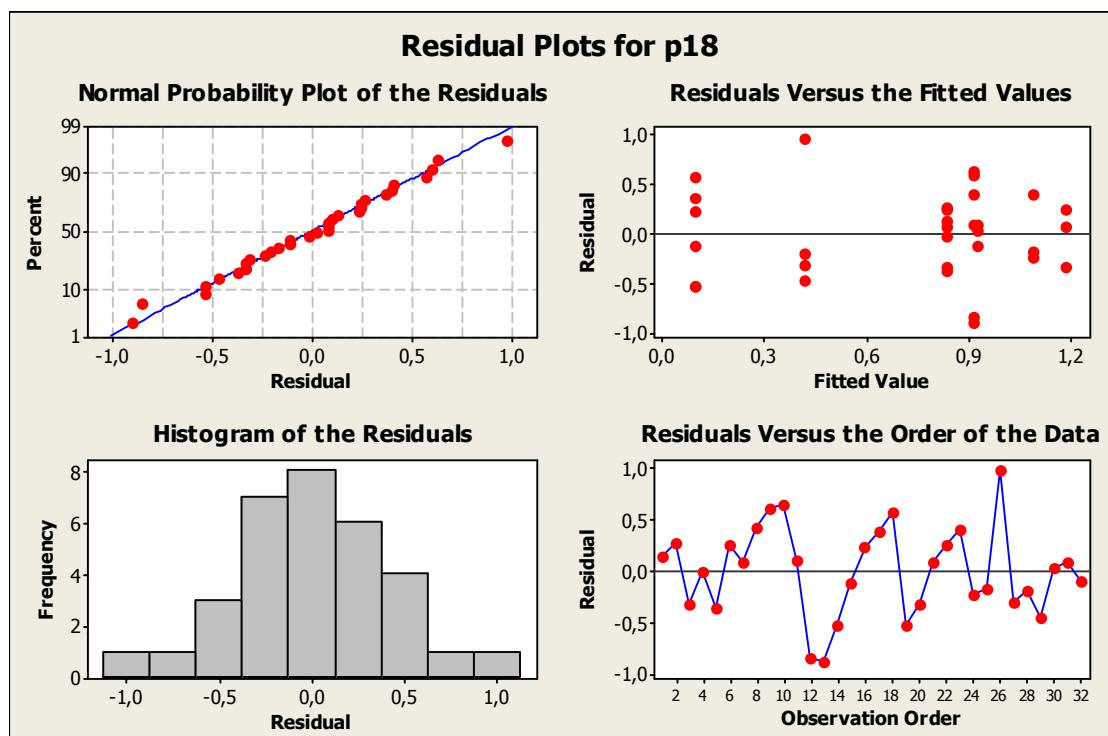


Gráfico 3.18 : Gráficos de Diagnóstico do Modelo

- ANOVA para o Metabólito P19 – Mínimos Quadrados Generalizados

Fonte	DF	Seq SS	Adj SS	Adj MS	F	P
Linhagem	1	3,345	1,153	1,153	1,04	0,318
Oxigênio	1	10,980	0,335	0,335	0,30	0,588
Glicose	1	2,212	0,267	0,267	0,24	0,628
Linhagem*Oxigênio	1	28,708	7,727	7,727	6,97	0,014
Linhagem*Glicose	1	37,053	12,872	12,872	11,61	0,002
Oxigênio*Glicose	1	32,693	32,693	32,693	29,49	0
Erro	25	27,717	27,717	1,109		
Total	31	142,709				

$$S = 1,05294 \quad R-Sq = 80,58\% \quad R-Sq(adj) = 75,92\%$$

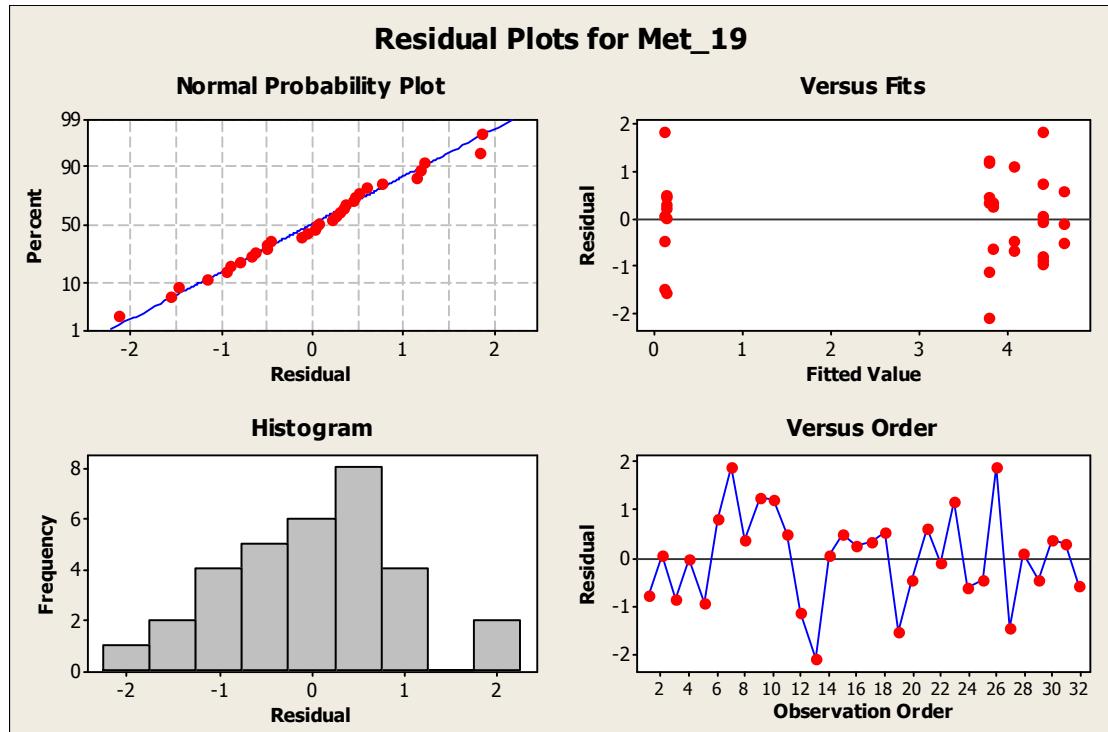


Gráfico 3.19 : Gráficos de Diagnóstico do Modelo

- ANOVA para o Metabólito P20

Fonte	DF	Seq SS	Adj SS	Adj MS	F	P
Linhagem	1	0,00963	0,04788	0,04788	0,52	0,477
Oxigênio	1	0,00193	0,01402	0,01402	0,15	0,699
Glicose	1	0,32498	0,36809	0,36809	4,01	0,056
Linhagem*Oxigênio	1	0,27	0,02966	0,02966	0,32	0,575
Linhagem*Glicose	1	0,01436	0,04754	0,04754	0,52	0,479
Oxigênio*Glicose	1	0,08543	0,08543	0,08543	0,93	0,344
Erro	25	2,29674	2,29674	0,09187		
Total	31	2,73333				

$$S = 0,303100 \quad R-Sq = 15,97\% \quad R-Sq(adj) = 0,00\%$$

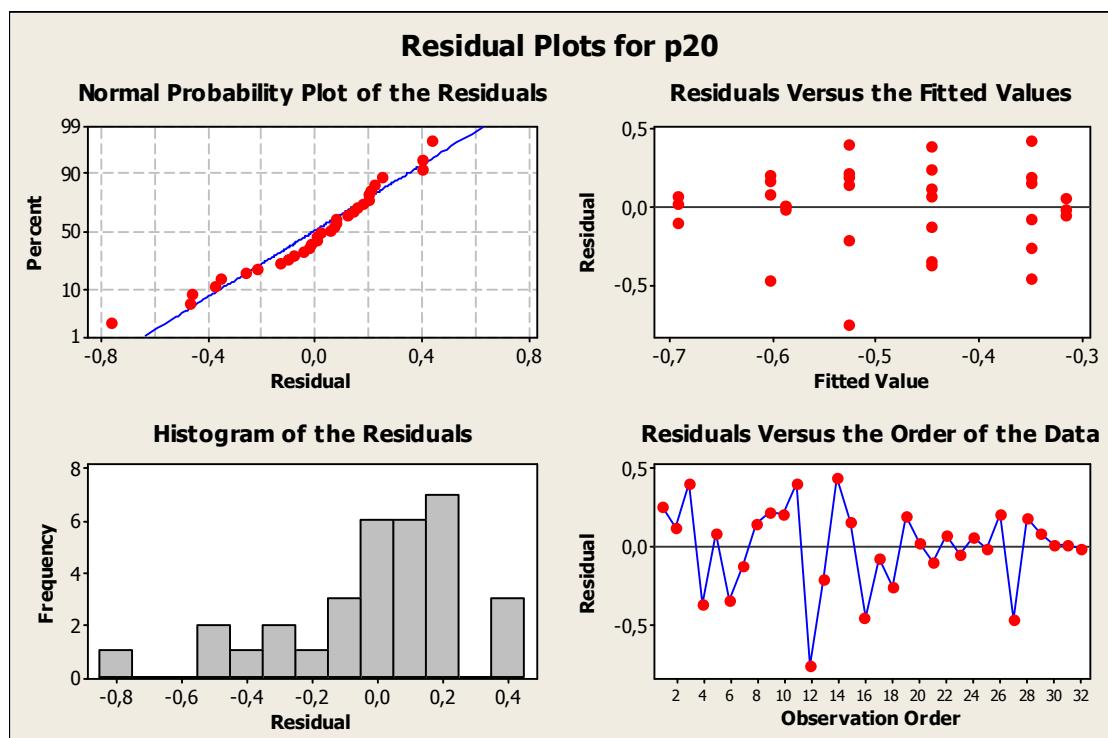


Gráfico 3.20 : Gráficos de Diagnóstico do Modelo

- ANOVA para o Metabólito P21

Fonte	DF	Seq SS	Adj SS	Adj MS	F	P
Linhagem	1	0,0015	0,1422	0,1422	1,28	0,269
Oxigênio	1	0,4180	0,5384	0,5384	4,83	0,037
Glicose	1	0,1155	0,3342	0,3342	3,00	0,096
Linhagem*Oxigênio	1	0,0032	0,0700	0,0700	0,63	0,436
Linhagem*Glicose	1	0,0444	0,1386	0,1386	1,24	0,275
Oxigênio*Glicose	1	0,2341	0,2341	0,2341	2,10	0,160
Erro	25	2,7853	2,7853	0,1114		
Total	31	3,6020				

$$S = 0,333787 \quad R-Sq = 22,67\% \quad R-Sq(adj) = 4,11\%$$

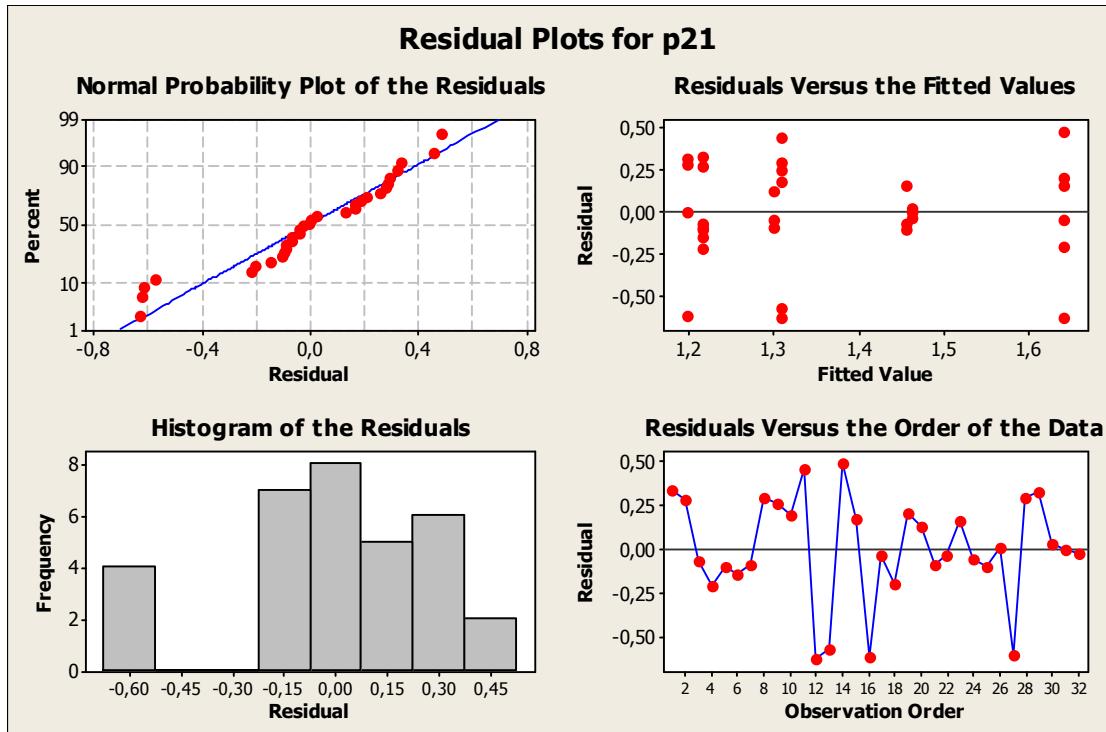


Gráfico 3.21 : Gráficos de Diagnóstico do Modelo

- ANOVA para o Metabólito P22

Fonte	DF	Seq SS	Adj SS	Adj MS	F	P
Linhagem	1	0,02218	0,01191	0,01191	0,17	0,684
Oxigênio	1	0,03424	0,00195	0,00195	0,03	0,869
Glicose	1	0,09307	0,17097	0,17097	2,44	0,131
Linhagem*Oxigênio	1	0,01321	0,03935	0,03935	0,56	0,460
Linhagem*Glicose	1	0,15538	0,13901	0,13901	1,99	0,171
Oxigênio*Glicose	1	0,25	0,25	0,25	0,00	0,953
Erro	25	1,75074	1,75074	0,07003		
Total	31	2,06907				

$$S = 0,264631 \quad R-Sq = 15,39\% \quad R-Sq(adj) = 0,00\%$$

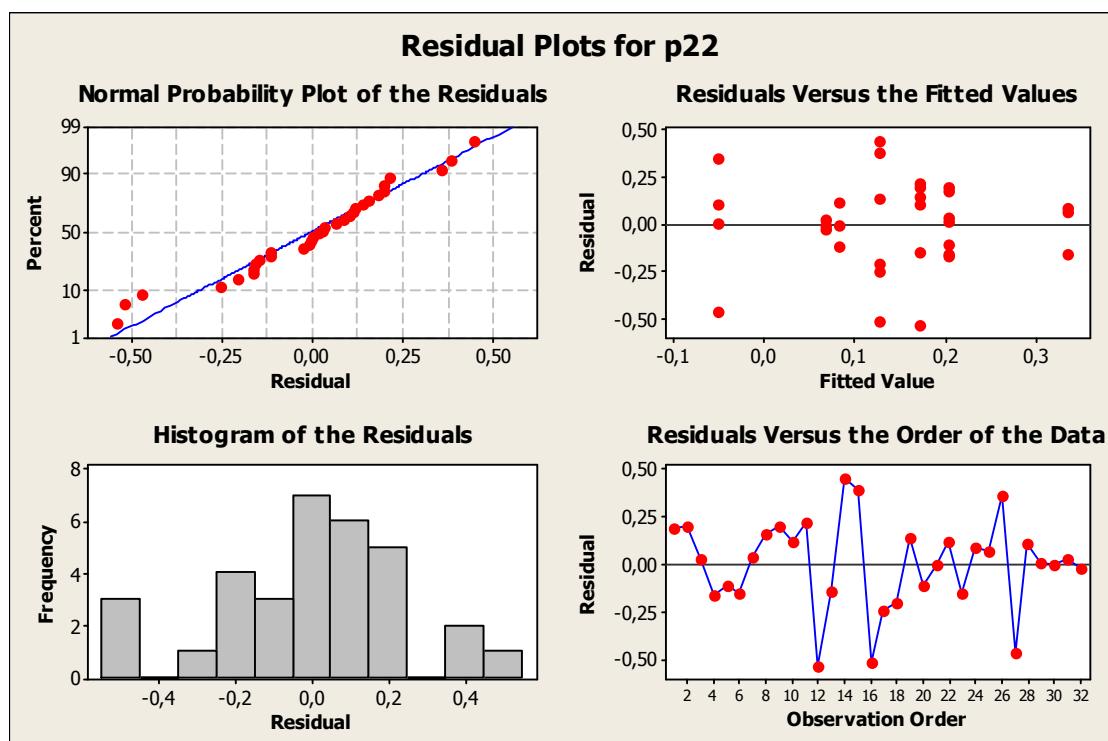


Gráfico 3.22 : Gráficos de Diagnóstico do Modelo

- ANOVA para o Metabólito P23

Fonte	DF	Seq SS	Adj SS	Adj MS	F	P
Linhagem	1	0,51778	0,58272	0,58272	11,25	0,003
Oxigênio	1	0,05066	0,01404	0,01404	0,27	0,607
Glicose	1	0,24635	0,10	0,10	0,00	0,966
Linhagem*Oxigênio	1	0,04884	0,24884	0,24884	4,80	0,038
Linhagem*Glicose	1	0	0,04717	0,04717	0,91	0,349
Oxigênio*Glicose	1	0,36018	0,36018	0,36018	6,95	0,014
Erro	25	1,29481	1,29481	0,05179		
Total	31	2,51861				

$$S = 0,227579 \quad R-Sq = 48,59\% \quad R-Sq(adj) = 36,25\%$$

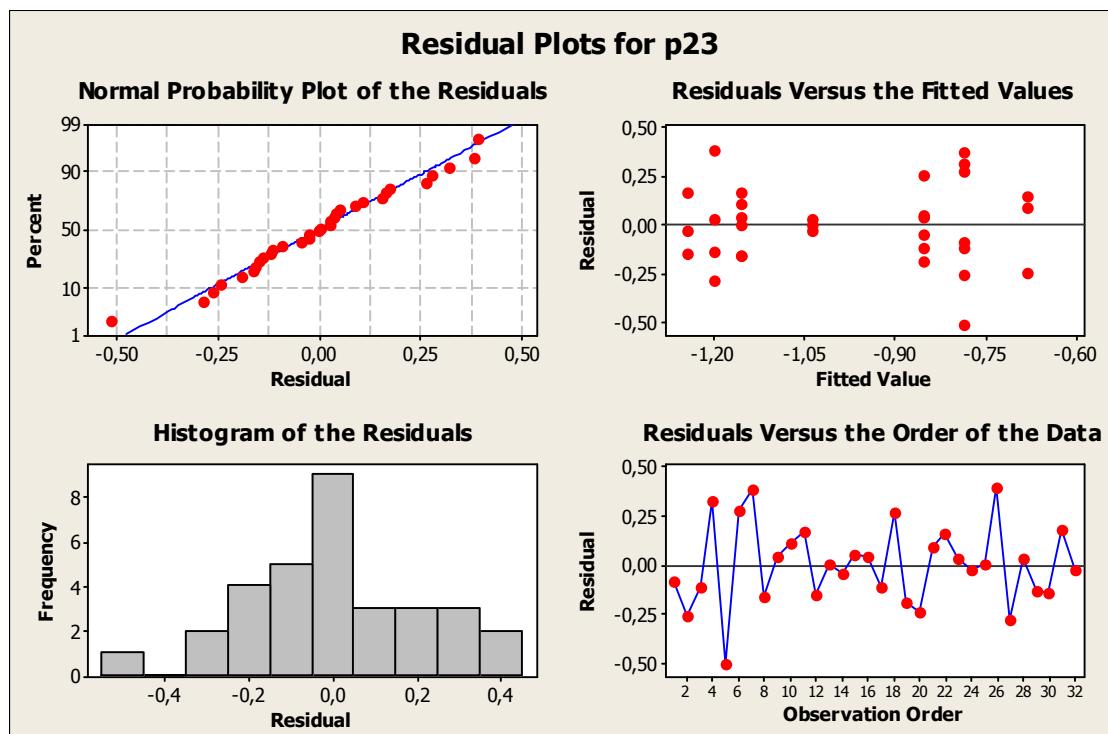


Gráfico 3.23 : Gráficos de Diagnóstico do Modelo

- ANOVA para o Metabólito P24

Fonte	DF	Seq SS	Adj SS	Adj MS	F	P
Linhagem	1	0	0,0618	0,0618	0,39	0,539
Oxigênio	1	1,3053	1,0185	1,0185	6,40	0,018
Glicose	1	0,0349	0,0931	0,0931	0,58	0,452
Linhagem*Oxigênio	1	0,0032	0,0052	0,0052	0,03	0,859
Linhagem*Glicose	1	0,1332	0,1286	0,1286	0,81	0,377
Oxigênio*Glicose	1	0,0026	0,0026	0,0026	0,02	0,899
Erro	25	3,9782	3,9782	0,1591		
Total	31	5,4574				

$$S = 0,398911 \quad R-Sq = 27,10\% \quad R-Sq(adj) = 9,61\%$$

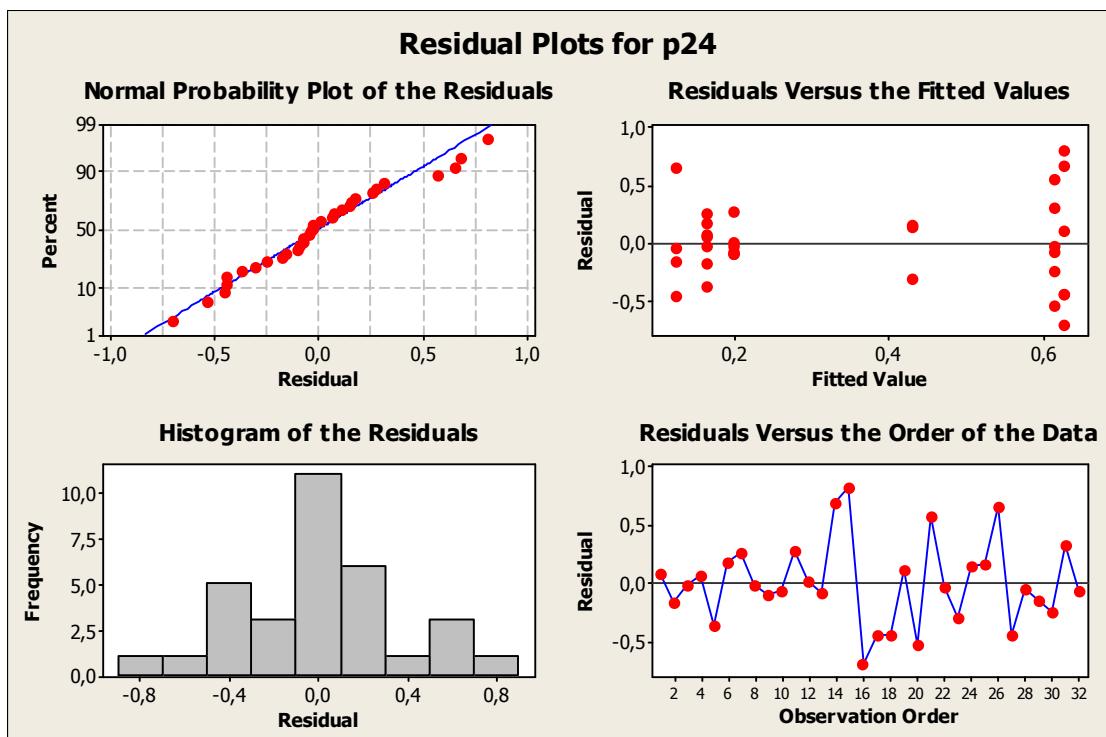


Gráfico 3.24 : Gráficos de Diagnóstico do Modelo

- ANOVA para o Metabólito P25

Fonte	DF	Seq SS	Adj SS	Adj MS	F	P
Linhagem	1	0,25247	0,48016	0,48016	5,65	0,025
Oxigênio	1	0,16749	0,49537	0,49537	5,83	0,023
Glicose	1	0,15829	0,01056	0,01056	0,12	0,728
Linhagem*Oxigênio	1	0,19032	0,33387	0,33387	3,93	0,059
Linhagem*Glicose	1	0,18767	0,23438	0,23438	2,76	0,109
Oxigênio*Glicose	1	0,04908	0,04908	0,04908	0,58	0,455
Erro	25	2,12596	2,12596	0,08504		
Total	31	3,13128				

$$S = 0,291613 \quad R-Sq = 32,11\% \quad R-Sq(adj) = 15,81\%$$

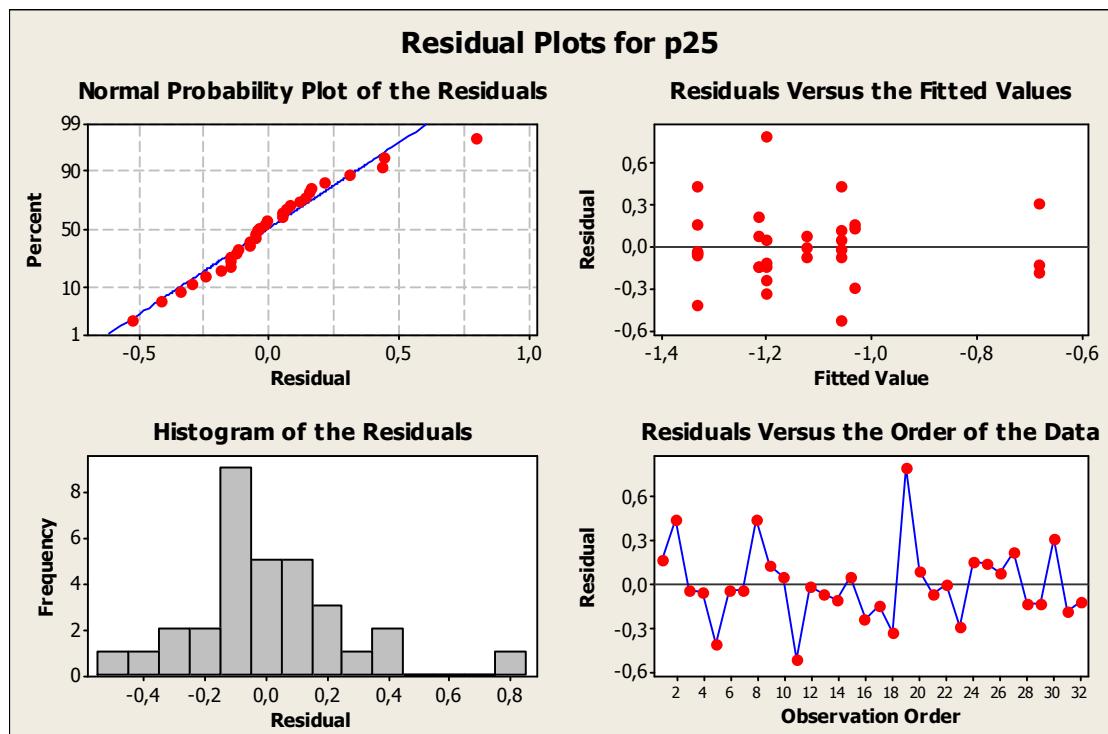


Gráfico 3.25 : Gráficos de Diagnóstico do Modelo

- ANOVA para o Metabólito P26

Fonte	DF	Seq SS	Adj SS	Adj MS	F	P
Linhagem	1	0,7817	0,0991	0,0991	0,82	0,373
Oxigênio	1	0,0598	0,0243	0,0243	0,20	0,657
Glicose	1	0,3451	0,4608	0,4608	3,83	0,062
Linhagem*Oxigênio	1	0,2656	0,0019	0,0019	0,02	0,900
Linhagem*Glicose	1	0,3165	0,0133	0,0133	0,11	0,742
Oxigênio*Glicose	1	1,2599	1,2599	1,2599	10,46	0,003
Erro	25	3,0106	3,0106	0,1204		
Total	31	6,0393				

$$S = 0,347023 \quad R-Sq = 50,15\% \quad R-Sq(adj) = 38,19\%$$

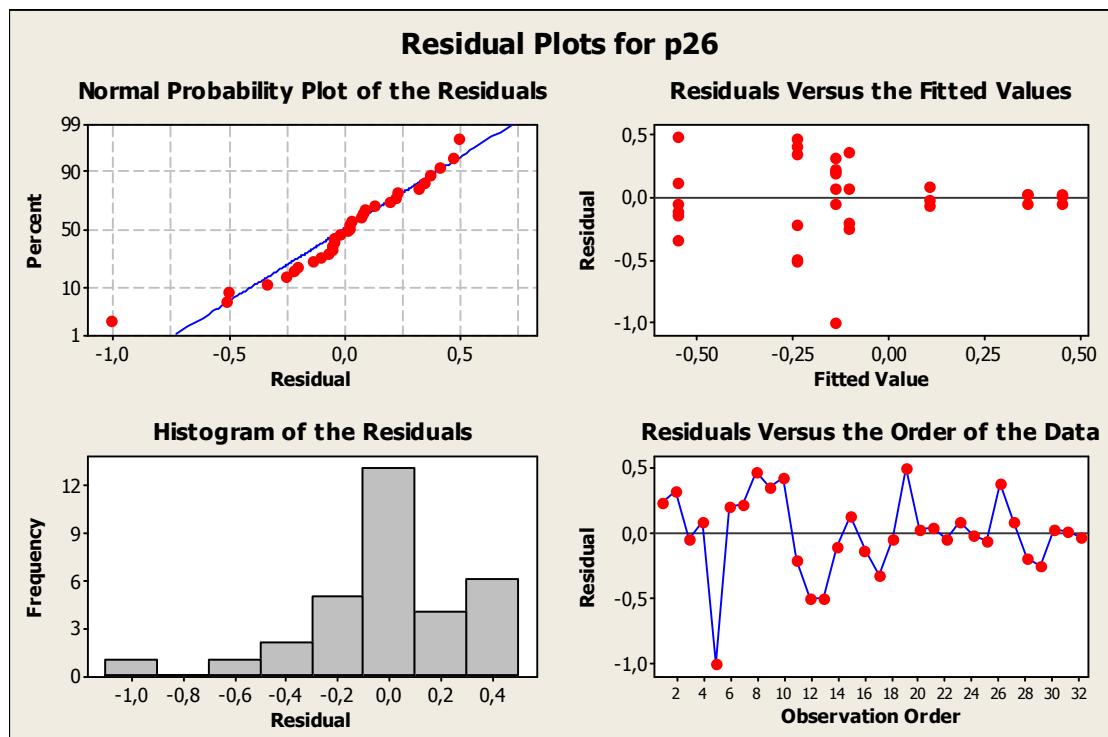


Gráfico 3.26 : Gráficos de Diagnóstico do Modelo

- ANOVA para o Metabólito P27

Fonte	DF	Seq SS	Adj SS	Adj MS	F	P
Linhagem	1	0,27280	0,11834	0,11834	1,74	0,200
Oxigênio	1	0,12837	0,33693	0,33693	4,94	0,036
Glicose	1	0,23635	1,01448	1,01448	14,88	0,001
Linhagem*Oxigênio	1	0,26599	0,20697	0,20697	3,03	0,094
Linhagem*Glicose	1	0,39038	0,01384	0,01384	0,20	0,656
Oxigênio*Glicose	1	3,69223	3,69223	3,69223	54,14	0
Erro	25	1,70499	1,70499	0,06820		
Total	31	6,69110				

$$S = 0,261150 \quad R-Sq = 74,52\% \quad R-Sq(adj) = 68,40\%$$

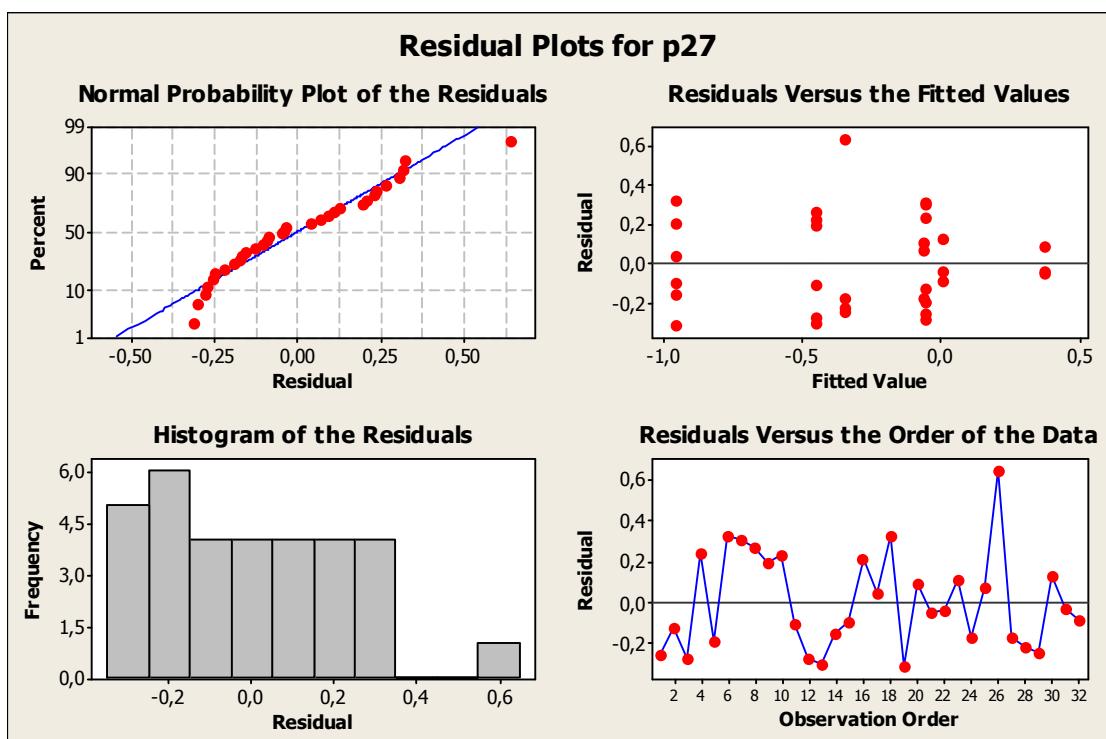


Gráfico 3.27 : Gráficos de Diagnóstico do Modelo

- ANOVA para o Metabólito P28 – Mínimos Quadrados Generalizados

Fonte	DF	Seq SS	Adj SS	Adj MS	F	P
Linhagem	1	2,7650	6,2954	6,2954	9,23	0,006
Oxigênio	1	1,0183	2,3639	2,3639	3,47	0,074
Glicose	1	0,0145	1,6233	1,6233	2,38	0,135
Linhagem*Oxigênio	1	0,2440	1,5199	1,5199	2,23	0,148
Linhagem*Glicose	1	1,0119	0,1191	0,1191	0,17	0,680
Oxigênio*Glicose	1	12,3994	12,3994	12,3994	18,18	0
Erro	25	17,0470	17,0470	0,6819		
Total	31	34,5001				

$$S = 0,825761 \quad R-Sq = 50,59\% \quad R-Sq(adj) = 38,73\%$$

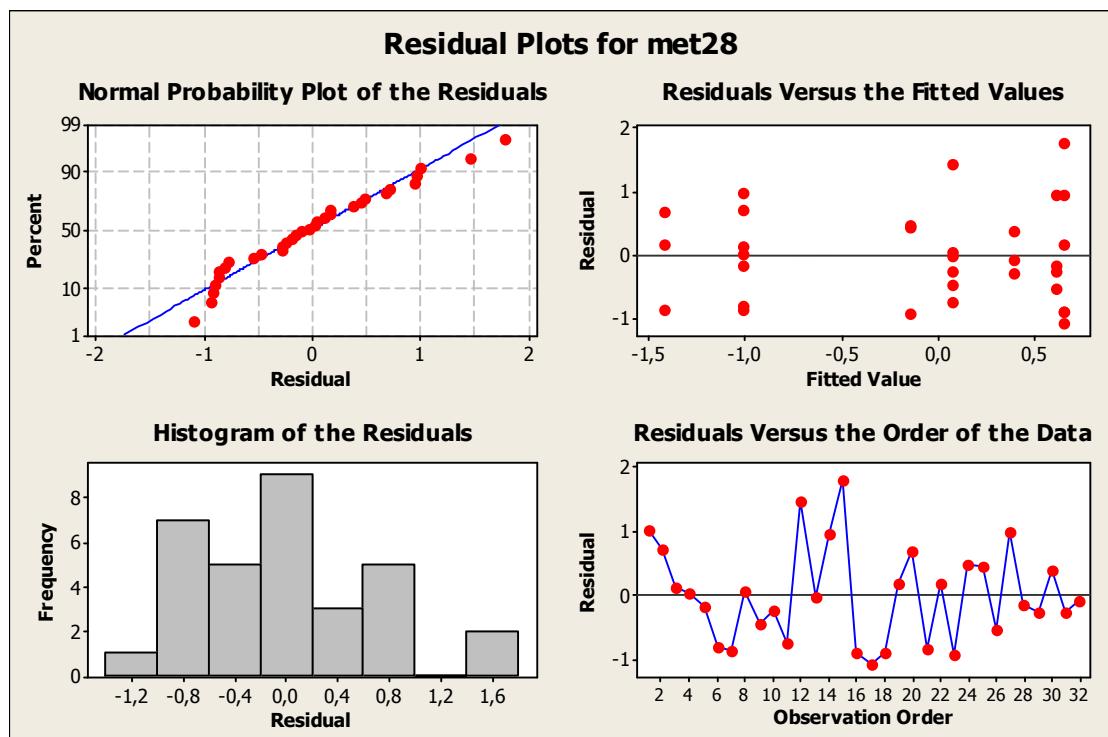


Gráfico 3.28 : Gráficos de Diagnóstico do Modelo

- ANOVA para o Metabólito P29

Fonte	DF	Seq SS	Adj SS	Adj MS	F	P
Linhagem	1	0,00589	0,01214	0,01214	0,26	0,614
Oxigênio	1	0,67559	0,28394	0,28394	6,10	0,021
Glicose	1	0,03053	0,00327	0,00327	0,07	0,793
Linhagem*Oxigênio	1	0,02769	0,00547	0,00547	0,12	0,735
Linhagem*Glicose	1	0,28063	0,13184	0,13184	2,83	0,105
Oxigênio*Glicose	1	0,12808	0,12808	0,12808	2,75	0,110
Erro	25	1,16373	1,16373	0,04655		
Total	31	2,31213				

$$S = 0,215752 \quad R-Sq = 49,67\% \quad R-Sq(adj) = 37,59\%$$

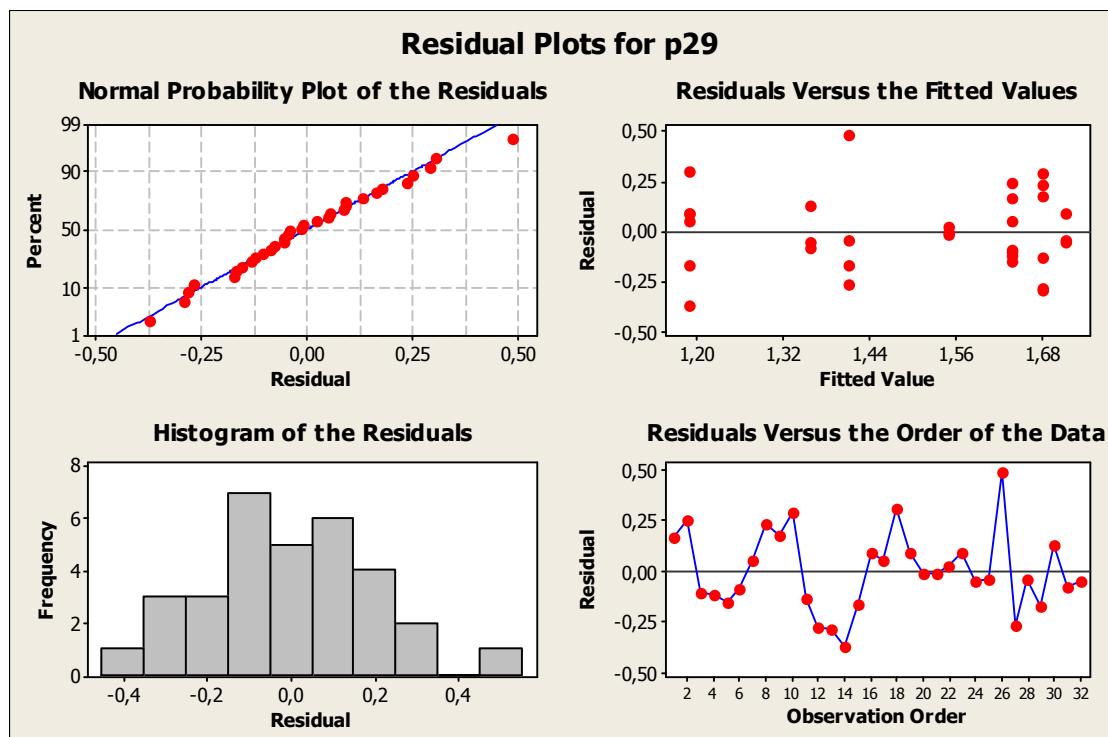


Gráfico 3.29 : Gráficos de Diagnóstico do Modelo

- ANOVA para o Metabólito P30 – Mínimos Quadrados Generalizados

Fonte	DF	Seq SS	Adj SS	Adj MS	F	P	
Linhagem	1	91,996	120,152	120,152	79,42	0	
Oxigênio	1	32,053	61,325	61,325	40,54	0	
Glicose	1	1,361	17,937	17,937	11,86	0,002	
Linhagem*Oxigênio	1	1,005	9,683	9,683	6,40	0,018	
Linhagem*Glicose	1	52,595	67,436	67,436	44,57	0	
Oxigênio*Glicose	1	16,024	1	6,024	16,024	10,59	0,003
Erro	25	37,822	37,822	1,513			
Total	31	232,855					

$$S = 1,22999 \quad R-Sq = 83,76\% \quad R-Sq(adj) = 79,86\%$$

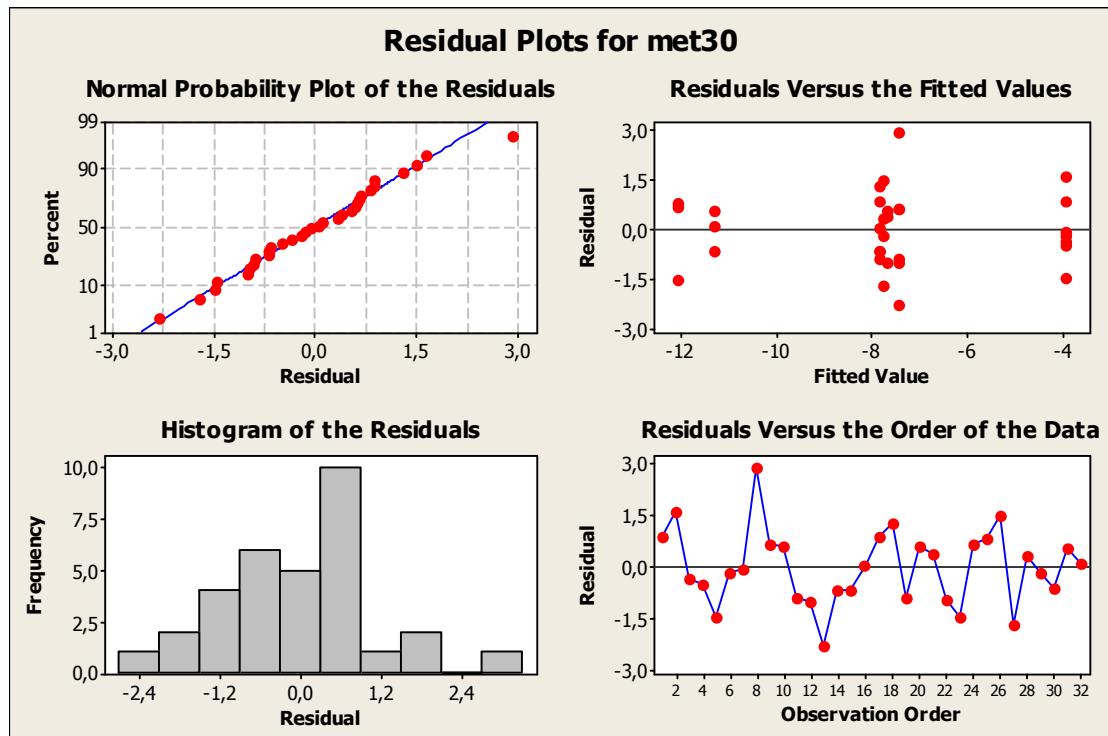


Gráfico 3.30 : Gráficos de Diagnóstico do Modelo

- ANOVA para o Metabólito P31

Fonte	DF	Seq SS	Adj SS	Adj MS	F	P
Linhagem	1	0,04076	0,07243	0,07243	1,02	0,323
Oxigênio	1	2,25600	1,51539	1,51539	21,25	0
Glicose	1	0,54442	0,49536	0,49536	6,95	0,014
Linhagem*Oxigênio	1	0,06607	0,00429	0,00429	0,06	0,808
Linhagem*Glicose	1	0,05678	0,02	0,02	0,00	0,987
Oxigênio*Glicose	1	0,38685	0,38685	0,38685	5,43	0,028
Erro	25	1,78258	1,78258	0,07130		
Total	31	5,13346				

$$S = 0,267026 \quad R-Sq = 65,28\% \quad R-Sq(adj) = 56,94\%$$

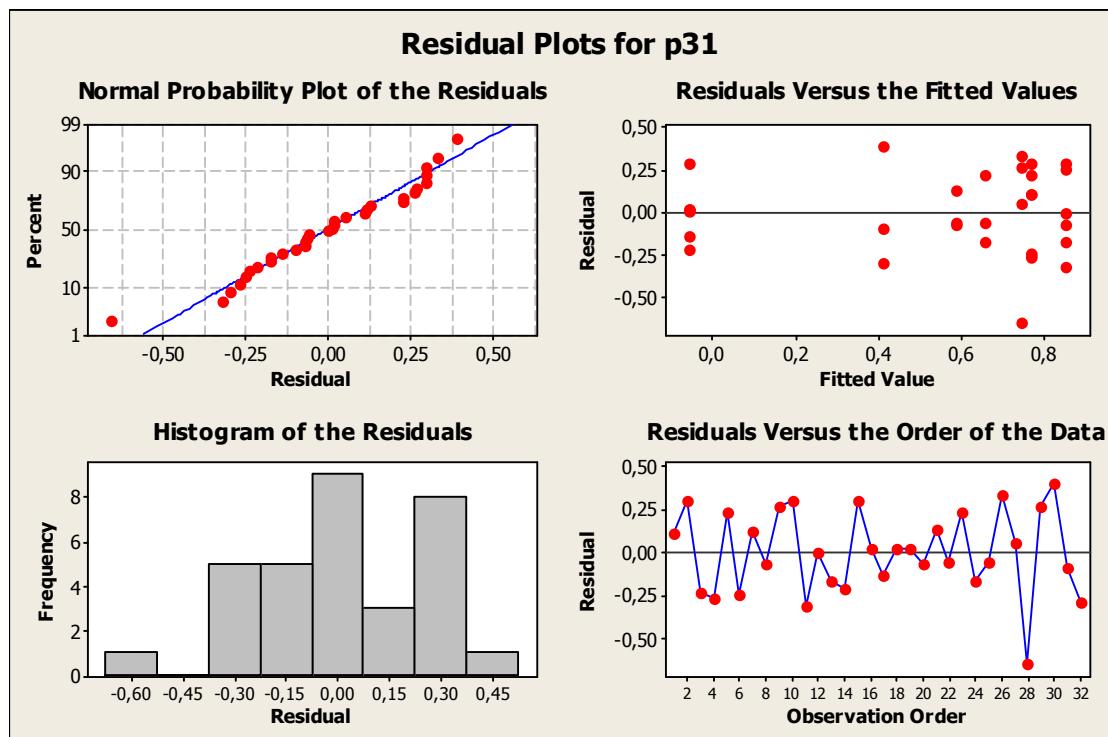


Gráfico 3.31 : Gráficos de Diagnóstico do Modelo

- ANOVA para o Metabólito Poutros

Fonte	DF	Seq SS	Adj SS	Adj MS	F	P
Linhagem	1	0,06618	0,00469	0,00469	0,23	0,638
Oxigênio	1	0,00447	0,01877	0,01877	0,91	0,350
Glicose	1	0,00468	0,01634	0,01634	0,79	0,383
Linhagem*Oxigênio	1	0,01384	0,02687	0,02687	1,30	0,265
Linhagem*Glicose	1	0,00347	0	0	0,00	0,999
Oxigênio*Glicose	1	0,02431	0,02431	0,02431	1,17	0,289
Erro	25	0,51765	0,51765	0,02071		
Total	31	0,63460				

$$S = 0,143895 \quad R-Sq = 18,43\% \quad R-Sq(adj) = 0,00\%$$

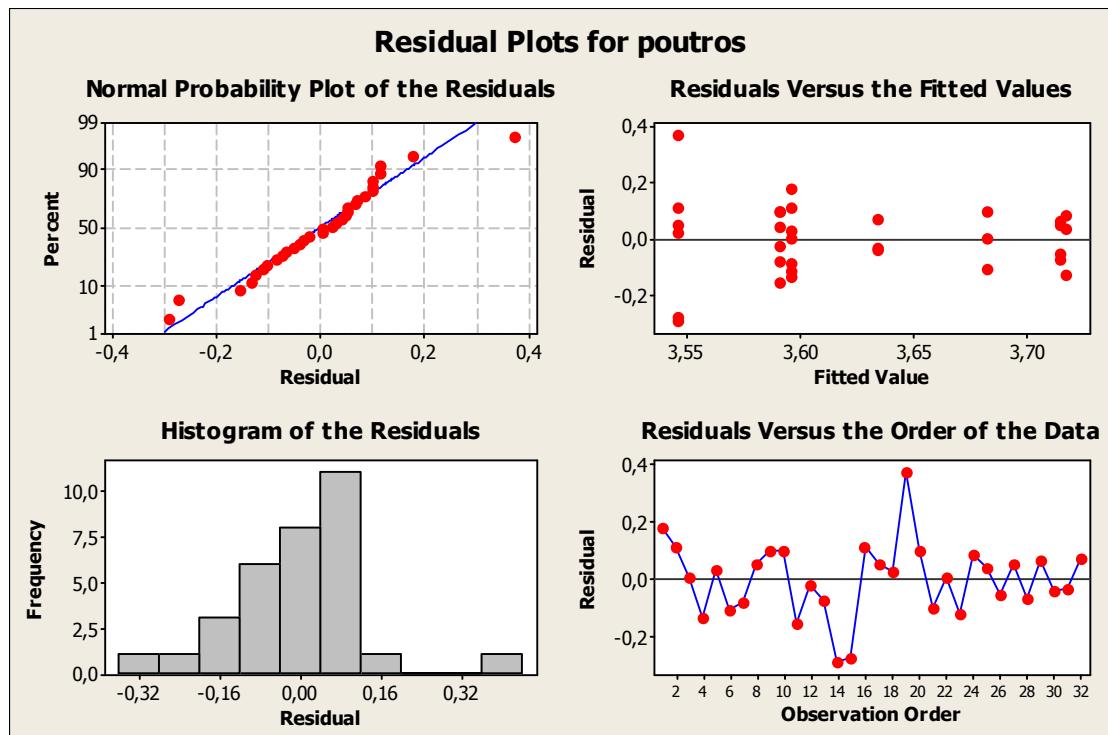


Gráfico 3.32 : Gráficos de Diagnóstico do Modelo

Apêndice 4: Análise Discriminante

Tabela 4.1 : Taxa de Acerto na Análise Discriminante

Fator	Taxa de Acerto	Cros-Validation
Glicose	84,40%	71,9%
Oxigênio	84,40%	78,10%
Linhagem	84,40%	84,40%

Funções Discriminantes:

1) Fator Glicose

Função Discriminante Linear para Grupos:

	Alta	Controle
Constante	-77,87	-70,97
P4	-4,04	1,45
P5	35,89	27,59
P12	-39,70	-45,33
P13	-6,40	-5,42
P17	-30,09	-26,90
P18	21,121	24,891
P27	-6,54	-3,30

2) Fator Oxigênio

Função Discriminante Linear para Grupos:

	Hipóxia	Normóxia
Constante	-17,851	-24,350
P6	-1,614	4,033
P24	4,968	3,671
P29	24,458	29,463

3) Fator Linhagem

Função Discriminante Linear para Grupos:

	G2	N3
Constant	-35,962	-25,752
P3	-13,275	-6,711
P5	47,433	38,865
P16	6,563	4,621

Apêndice 5: Regressão Logística

Modelo Logístico para o Fator Linhagem

Modelo 5.1: Regressão Logística para o Fator Linhagem

Analysis of Maximum Likelihood Estimates					
Parameter	DF	Estimate	Standard Error	Chi-Square	Pr > ChiSq
Intercept	1	-8.5594	4.2732	4.0121	0.0452
p3	1	-6.2894	2.6169	5.7762	0.0162
p5	1	8.0275	3.3819	5.6344	0.0176
p16	1	1.9884	0.8476	5.5029	0.0190

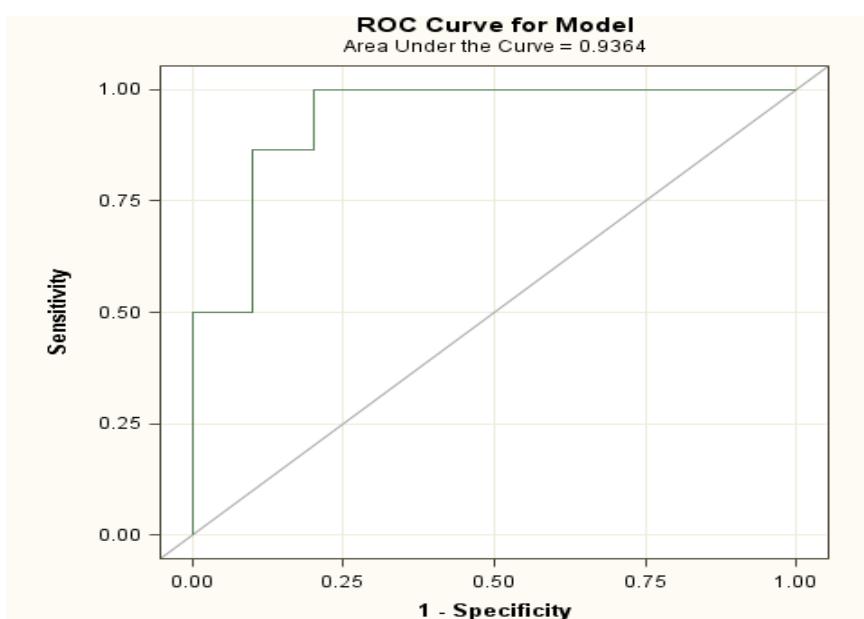


Gráfico 4.1: Curva Roc para o Modelo 5.1

Tabela 5.1 : Tabela de Classificação para o Modelo 5.1

<u>FROM</u>	<u>INTO</u>	IP_G2	IP_N3
G2	G2	0.96258	0.03742
G2	G2	0.89969	0.10031
G2	G2	0.99729	0.00271
G2	G2	0.9785	0.0215
G2	G2	0.99389	0.00611
G2	G2	0.93236	0.06764
G2	G2	0.99003	0.00997
G2	N3	0.33698	0.66302
G2	N3	0.42978	0.57022
G2	G2	0.809	0.191
G2	G2	0.99471	0.00529
G2	G2	0.81443	0.18557
G2	G2	0.80671	0.19329
G2	G2	0.93747	0.06253
G2	G2	0.99905	0.00095
G2	G2	0.98554	0.01446
G2	G2	0.93372	0.06628
G2	G2	0.99372	0.00628
G2	G2	0.50494	0.49506
G2	G2	0.96403	0.03597
G2	G2	0.98282	0.01718
G2	G2	0.84119	0.15881
N3	G2	0.95305	0.04695
N3	N3	0.03902	0.96098
N3	N3	0.01325	0.98675
N3	G2	0.69676	0.30324
N3	N3	0.22855	0.77145
N3	N3	0.29825	0.70175
N3	N3	0.28763	0.71237
N3	N3	0.1716	0.8284
N3	N3	0.20257	0.79743
N3	N3	0.02089	0.97911

FROM : Representa a classificação real dos Fatores de Controle.

INTO : Representa a classificação do Modelo Logístico

IP_G2 : probabilidade de pertencer ao nível G2 do Fator Linhagem

IP_N3 : probabilidade de pertencer ao nível N3 do Fator Linhagem

Observações:

- O nível G2 foi selecionado como nível de interesse.
- As observações em destaque foram classificadas erroneamente

Tabela 5.2 : Teste de Ajuste

Deviance and Pearson Goodness-of-Fit Statistics				
Criterion	Value	DF	Value/DF	Pr > ChiSq
Deviance	19.1655	28	0.6845	0.8930
Pearson	29.9075	28	1.0681	0.3677

Modelo Logístico para o Fator Oxigênio

Modelo 5.2: Regressão Logística para o Fator Oxigênio

Analysis of Maximum Likelihood Estimates					
Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	12.6308	7.6887	2.6987	0.1004
p6	1	-7.4370	4.4938	2.7389	0.0979
p21	1	1.4580	1.8990	0.5894	0.4426
p24	1	3.2513	2.8378	1.3127	0.2519
p29	1	-10.8984	5.2615	4.2904	0.0383

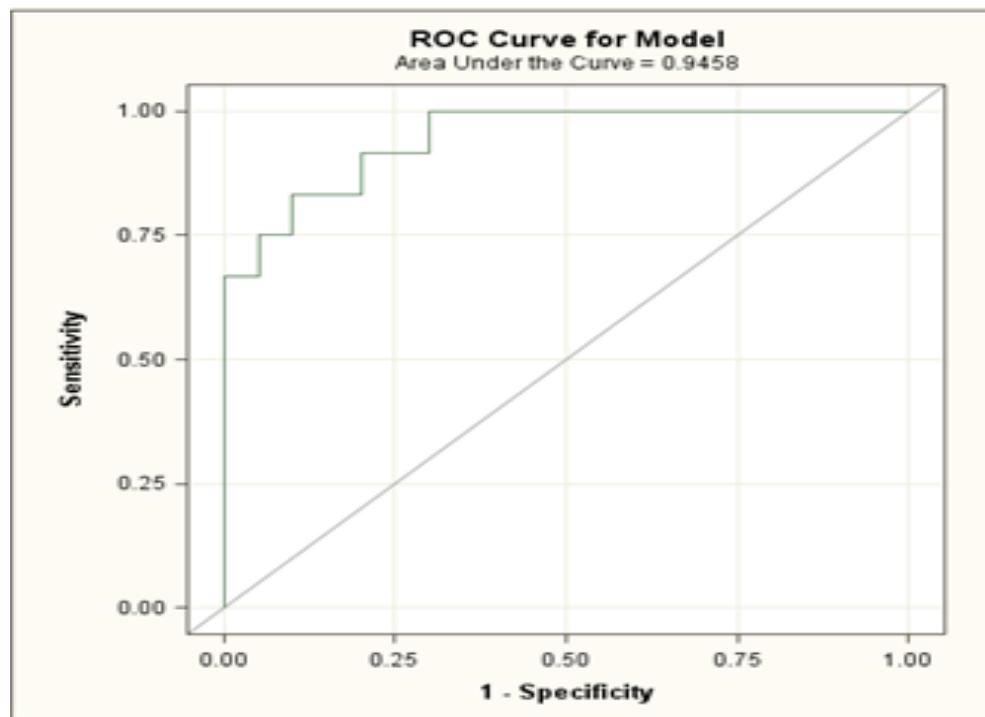


Gráfico 5.2: Curva Roc para o Modelo 5.2

Tabela 5.3 : Tabela de Classificação para o Modelo 5.2

FROM	_INTO_	IP_Hip_xia	IP_Norm_xia
Normóxia	Normóxia	0.00558	0.99442
Normóxia	Normóxia	0.0011	0.9989
Normóxia	Normóxia	0.07215	0.92785
Normóxia	Normóxia	0.07714	0.92286
Normóxia	Normóxia	0.05615	0.94385
Normóxia	Normóxia	0.0252	0.9748
Normóxia	Normóxia	0.00656	0.99344
Normóxia	Normóxia	0.0092	0.9908
Normóxia	Normóxia	0.01037	0.98963
Normóxia	Normóxia	0.00312	0.99688
Normóxia	Hipóxia	0.53805	0.46195
Normóxia	Normóxia	0.1895	0.8105
Normóxia	Normóxia	0.45696	0.54304
Hipóxia	Hipóxia	1	0
Hipóxia	Hipóxia	1	0
Hipóxia	Normóxia	0.19838	0.80162
Hipóxia	Hipóxia	0.78332	0.21668
Hipóxia	Normóxia	0.16635	0.83365
Hipóxia	Hipóxia	0.98955	0.01045
Hipóxia	Hipóxia	0.53583	0.46417
Hipóxia	Hipóxia	0.97175	0.02825
Hipóxia	Hipóxia	0.56933	0.43067
Normóxia	Normóxia	0.00362	0.99638
Normóxia	Normóxia	0.17227	0.82773
Normóxia	Normóxia	0.13097	0.86903
Normóxia	Normóxia	0.00948	0.99052
Normóxia	Normóxia	0.07315	0.92685
Normóxia	Hipóxia	0.63412	0.36588
Normóxia	Hipóxia	0.51964	0.48036
Hipóxia	Hipóxia	0.96424	0.03576
Hipóxia	Hipóxia	0.99895	0.00105
Hipóxia	Hipóxia	0.82798	0.17202

FROM : Representa a classificação real dos Fatores de Controle.

INTO : Representa a classificação do Modelo Logístico

IP_Hip_xia : probabilidade de pertencer ao nível Hipóxia do Fator Oxigênio

IP_Norm_xia : probabilidade de pertencer ao nível Normóxia do Fator Oxigênio

Observações:

- O nível Hipóxia foi selecionado como nível de interesse.
- As observações em destaque foram classificadas erroneamente

Tabela 5.4 : Teste de Ajuste

Deviance and Pearson Goodness-of-Fit Statistics				
Criterion	Value	DF	Value/DF	Pr > ChiSq
Deviance	18.2656	27	0.6765	0.8952
Pearson	17.0257	27	0.6306	0.9305

Modelo Logístico para o Fator Glicose

Modelo 5.3: Regressão Logística para o Fator Glicose

Analysis of Maximum Likelihood Estimates					
Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	6.5567	5.5303	1.4057	0.2358
p4	1	-6.1623	2.7882	4.8848	0.0271
p12	1	3.2905	3.6382	0.8180	0.3658
p16	1	0.5429	0.5874	0.8541	0.3554
p18	1	-3.0880	2.1584	2.0469	0.1525
p27	1	-1.7923	1.7276	1.0763	0.2995
p31	1	-1.7110	1.3374	1.6367	0.2008

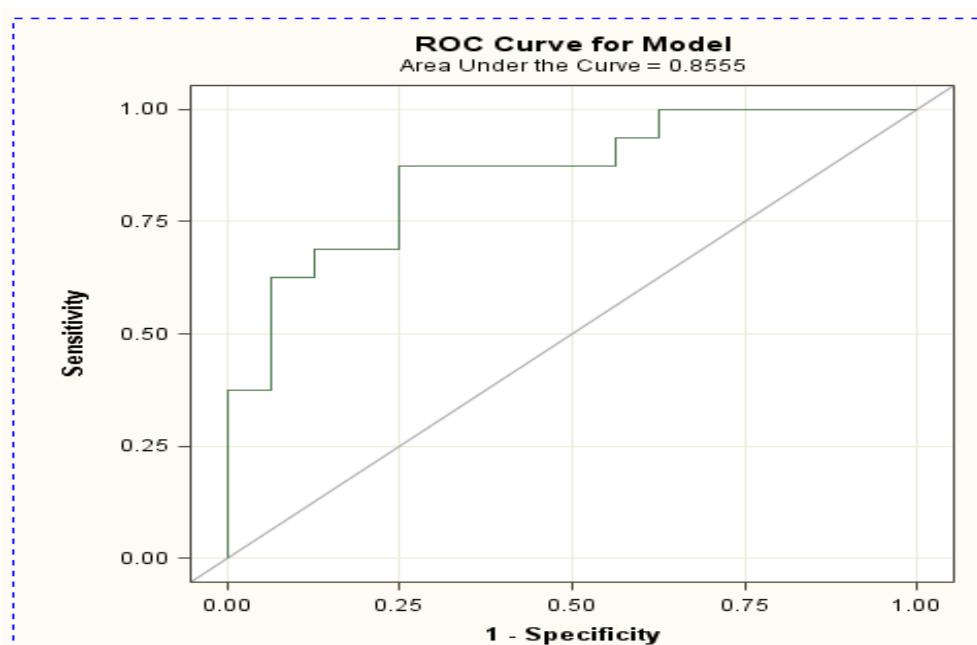


Gráfico 5.3 : Curva Roc para o Modelo 5.3

Tabela 5.5: Tabela de Classificação para o Modelo 5.3

<u>FROM</u>	<u>INTO</u>	IP_Alta	IP_Ctrl
Alta	Alta	0.77561	0.22439
Alta	Ctrl	0.46179	0.53821
Alta	Alta	0.95472	0.04528
Alta	Alta	0.6589	0.3411
Alta	Ctrl	0.16429	0.83571
Alta	Alta	0.71349	0.28651
Alta	Alta	0.90391	0.09609
Ctrl	Ctrl	0.08885	0.91115
Ctrl	Ctrl	0.31347	0.68653
Ctrl	Ctrl	0.26831	0.73169
Ctrl	Alta	0.665	0.335
Ctrl	Ctrl	0.09416	0.90584
Ctrl	Ctrl	0.1258	0.8742
Alta	Alta	0.98517	0.01483
Alta	Alta	0.98357	0.01643
Alta	Ctrl	0.45918	0.54082
Alta	Alta	0.83587	0.16413
Alta	Alta	0.85896	0.14104
Alta	Alta	0.99279	0.00721
Ctrl	Ctrl	0.34456	0.65544
Ctrl	Ctrl	0.49219	0.50781
Ctrl	Ctrl	0.24085	0.75915
Alta	Ctrl	0.24091	0.75909
Alta	Alta	0.73487	0.26513
Alta	Ctrl	0.42653	0.57347
Ctrl	Alta	0.5248	0.4752
Ctrl	Ctrl	0.12472	0.87528
Ctrl	Alta	0.85885	0.14115
Ctrl	Ctrl	0.40642	0.59358
Ctrl	Ctrl	0.01958	0.98042
Ctrl	Ctrl	0.03423	0.96577
Ctrl	Ctrl	0.24764	0.75236

FROM : Representa a classificação real dos Fatores de Controle.

INTO : Representa a classificação do Modelo Logístico

IP_Alta : probabilidade de pertencer ao nível Alta do Fator Glicose

IP_Ctrl : probabilidade de pertencer ao nível Controle do Fator Glicose

Observações :

- O nível Alta foi selecionado como nível de interesse.
- As observações em destaque foram classificadas erroneamente.

Tabela 5.6 : Teste de Ajuste

Deviance and Pearson Goodness-of-Fit Statistics				
Criterion	Value	DF	Value/DF	Pr > ChiSq
Deviance	29.2938	25	1.1718	0.2518
Pearson	27.4149	25	1.0966	0.3355