

ADRIANO DE MORAES ARANTES

**PREDIÇÃO DE DOENÇA DO ENXERTO CONTRA HOSPEDEIRO
AGUDA BASEADA NO PERFIL DE EXPRESSÃO GÊNICA**

Tese apresentada à Universidade Federal
de São Paulo para obtenção do título de
Doutor em Ciências

SÃO PAULO
2009

ADRIANO DE MORAES ARANTES

**PREDIÇÃO DE DOENÇA DO ENXERTO CONTRA HOSPEDEIRO
AGUDA BASEADA NO PERFIL DE EXPRESSÃO GÊNICA**

Tese apresentada à Universidade Federal
de São Paulo para obtenção do título de
Doutor em Ciências

Orientador: Prof. Dr. José Salvador Rodrigues de Oliveira

Co-orientadores: Dra. Natalia Shulzhenko

Dr. Andrey Morgun

São Paulo

2009

Arantes, Adriano de Moraes

Predição de doença do enxerto contra hospedeiro aguda baseada no perfil de expressão gênica. Estudo Prospectivo/ Adriano de Moraes Arantes – São Paulo, 2009.

xviii, 155f

Tese (Doutorado) – Universidade Federal de São Paulo. Escola Paulista de Medicina. Programa de Pós-Graduação em Hematologia.

Título em Inglês: Acute graft versus host disease prediction based on gene expression profiling.

1. transplante alogênico de células tronco hematopoéticas. 2. doença do enxerto contra o hospedeiro aguda. 3. microarray. 4. perfil de expressão gênica.

*Aos meus pais, pelo apoio
irrestrito e carinho
contínuo.*

*À Viviane, pelo amor
incondicional e
companheirismo em todos
os momentos.*

*Às minhas queridas
irmãs Janaína e Flávia,
pelo carinho, apoio e
amizade.*

Agradecimentos

Ao Prof. Dr. José Salvador Rodrigues de Oliveira, por ter me recebido, pela oportunidade, pela orientação científica e pelos profundos ensinamentos em transplante de medula óssea.

À Profa. Dra. Maria Gerbase DeLima, pelos primeiros ensinamentos em metodologia científica, pelo incentivo e contribuições apresentadas.

Aos Drs Andrey Morgun e Natasha Shulzhenko, por acreditarem em meu potencial, por me criticarem nos momentos necessários, pela amizade irrestrita e pela generosidade em partilhar seus conhecimentos.

À Dra Chloé Musseti pelos constantes ensinamentos de Imunologia básica e o interesse contínuo em manter-nos atualizados.

Às Dras Gerdine Sanson e Susana Diniz, pela competência e forma crítica com que atuaram durante minha pós graduação. Pela grande amizade e respeito construída entre nós durante esta profícua convivência.

Ao Amador pela grande parceria, pela amizade, pela confiança e pela enorme dedicação em todos os momentos desse trabalho.

Ao Prof. Dr. José Orlando Bordin, pelo estímulo à pesquisa, pelo exemplo e pelo apoio.

À Profa Dra. Maria Stella Figueiredo, pelos conselhos, pela atenciosa ajuda e pelo apoio.

À Profa. Dra. Maria de Lourdes L. Ferrari Chauffaille, pela sinceridade, pela confiança e pelos ensinamentos.

Ao Prof. Dr. José Kerbauy pelo acolhimento nessa instituição, pelo incentivo e pelos ensinamentos.

Ao Celso Arrais Rodrigues, pela fiel amizade e lealdade, sempre disponível para os momentos de incertezas.

Ao Fábio Rodrigues Kerbauy , pela grande amizade construída e pela agradável convivência.

Aos amigos José Dias e Bruna Tirapelli , pela amizade e parceria contínua.

Aos amigos pós-graduandos, residentes, médicos e funcionários da Disciplina de Hematologia e Hemoterapia da UNIFESP, pela convivência agradável.

Aos amigos biólogos, plantonistas e funcionários do Laboratório de HLA da UNIFESP, minha gratidão por ensinar-me os primeiros passos em uma bancada de laboratório.

Esse trabalho foi desenvolvido com apoio financeiro da Associação Fundo de Incentivo à Psicofarmacologia (AFIP), CAPES e do “Ghost Laboratory”- National Institute of Allergy and Infectious Disease, National Institute of Health-EUA.

Sumário

Dedicatória-----	iii
Agradecimentos-----	vi
Lista de figuras-----	xi
Lista de tabelas-----	xii
Lista de abreviaturas-----	xiii
Resumo-----	xvi
1. INTRODUÇÃO-----	1
1.1. Descrição do problema: maior incidência de aGVHD moderada a grave implica em maior morbi-mortalidade devido a lesão de órgão alvo, infecções e outras complicações relacionadas ao GVHD-----	2
1.2. Uso de <i>microarrays</i> como ferramenta investigativa em hematologia e TCTH -----	4
1.3. Predição em GVHD agudo -----	7
2. OBJETIVOS-----	10
3. CASUÍSTICA E MÉTODOS-----	12
3.1. Casuística-----	13
3.2. Técnicas de transplantes utilizadas-----	14
3.3. Coleta de amostras e processamento-----	15
3.3.1. Processamento do material-----	15
3.3.2. Extração de RNA total -----	16
3.3.3. Avaliação da qualidade e quantidade do RNA total -----	16
3.3.4. Amplificação do RNA mensageiro (mRNA)-----	16

3.4. Método de <i>Microarray</i> -----	19
3.4.1. Escaneamento das Lâminas de <i>Microarray</i> -----	19
3.4.2. Filtragem dos dados gerados -----	22
3.4.3. Normalização da expressão gênica-----	22
3.5. Métodos estatísticos-----	27
3.5.1. Métodos estatísticos para análise de dados clínicos-----	27
3.5.2. Métodos estatísticos para geração e validação do classificador molecular-----	27
3.5.3. Análise de genes diferencialmente expressos-----	30
3.6. Análise funcional dos genes: a busca em bancos de dados-----	31
4. RESULTADOS -----	32
4.1. Características dos pacientes-----	33
4.2. Impacto da ocorrência de aGVHD e fatores que interferiram na sobrevida geral ao final do primeiro ano-----	36
4.3. Identificação dos genes que compõem o preditor de aGVHD e teste do preditor---	41
4.3.1. Identificação dos genes do preditor-----	41
4.3.2. Teste do preditor-----	42
4.4. Análise de genes diferencialmente expressos -----	45
4.5. Anotação funcional dos genes diferencialmente expressos -----	47
4.6. Agrupamento dos genes diferencialmente expressos segundo a expressão em diferentes tipos celulares e tecidos humanos-----	56
5. DISCUSSÃO -----	60
5.1. Predição de GVHD agudo como ferramenta diagnóstica-----	61

5.2. Microarrays e a descoberta de novos eventos biológicos envolvendo aGVHD----	65
6. CONCLUSÕES-----	76
7. ANEXOS-----	78
8. REFERÊNCIAS-----	138

Abstract

Bibliografia consultada

Lista de Figuras

Figura 1	Ilustração do método utilizado para amplificação do mRNA-----	18
Figura 2	Representação esquemática de um experimento de <i>microarray</i> -----	20
Figura 3	Exemplo de imagem de uma lâmina de <i>microarray</i> após escaneamento-----	21
Figura 4	Gráfico do tipo box-plot mostrando a intensidade do sinal de oito lâminas antes e após o procedimento de normalização -----	25
Figura 5	Gráfico do tipo MA plot de três lâminas antes e após o procedimento de normalização----	26
Figura 6	Representação esquemática da seleção de genes para o preditor e teste do preditor-----	29
Figura 7	Distribuição dos 89 pacientes incluídos de acordo com o hospital e tipo de doença-----	35
Figura 8	Gráfico de sobrevida global dos pacientes estratificados segundo classificação de aGVHD----- -----	39
Figura 9	Classificação de amostras de aGVHD e sem aGVHD nos grupos treino e teste segundo acertos em cada um dos sete algoritmos de predição-----	44
Figura 10	Heatmap dos genes diferencialmente expressos entre amostras dos pacientes que evoluíram e não evoluíram para aGVHD-----	46
Figura 11	Rede de comunicação bioquímica da via NF-kB/p38 MAPK entre genes sub- expressos----	55
Figura 12	Heatmap dos 54 genes super-expressos e o agrupamento dos mesmo em relação à expressão destes genes em diversos tecidos e linhagens celulares humanas obtidas do SymAtlas----	58
Figura 13	Heatmap dos 16 genes sub-expressos e o agrupamento dos mesmo em relação a expressão destes genes em diversos tecidos e linhagens celulares humanas obtidas do SymAtlas-----	59

Lista de Tabelas

Tabela 1	Características clínicas dos pacientes.-----	38
Tabela 2	Análise univariada e multivariada de Cox para a sobrevida geral ao final do primeiro ano de seguimento -----	40
Tabela 3	Desempenho do preditor nos grupos treino e teste após exclusão de amostras inclassificáveis-----	44
Tabela 4	Análise funcional do grupo genes super-expressos segundo envolvimento em processo biológico baseado no Gene Ontology (GO)-----	50
Tabela 5	Análise funcional do grupo genes super-expressos segundo sua função molecular baseado no Gene Ontology (GO)-----	51
Tabela 6	Análise funcional do grupo genes sub-expressos segundo envolvimento em processo biológico baseado no Gene Ontology (GO)-----	52
Tabela 7	Análise funcional do grupo genes sub-expressos segundo sua função molecular baseado no Gene Ontology (GO)-----	54

Lista de Abreviaturas

aGVHD	doença do enxerto contra o hospedeiro aguda
Aminoallyl-dUTP	(5-[3-aminoallyl]-2'-deoxyuridine -5'-triphosphate)
aRNA	RNA amplificado
<i>BPI</i>	bactericidal/permeability-increasing protein
CD4	antígeno CD4
CD8	antígeno CD8
cDNA	DNA complementar
cGVHD	doença do enxerto contra o hospedeiro crônica
CMV	citomegalovírus
CSA	ciclosporina
<i>CTSG</i>	Catepsina G
Cy3	corante Cy3
Cy5	corante Cy5
<i>DEFA4</i>	defensina A 4
DNA	ácido desoxirribonucléico
<i>ELA2</i>	elastase 2
<i>FasL</i>	Fas ligante
FDR	<i>false discovery rate – taxa de descobertas falsas</i>
G	guanina
GO	<i>Gene Ontology</i>

HLA	antígeno maior de histocompatibilidade
<i>IFI27</i>	interferon alpha-inducible protein 27
<i>IFNγ</i>	interferon gama
<i>IL-1α</i>	interleucina 1 alfa
<i>IL-1β</i>	interleucina 1 beta
<i>IL-12p40</i>	subunidade p40 da Interleucina 12
<i>IL-1Ra</i>	receptor antagonista da interleucina 1
<i>IL-2</i>	Interleucina 2
<i>IL-8</i>	Interleucina 8
<i>Lowess</i>	<i>Locally weighted linear regression</i>
LPS	lipopolissacáride
MMF	micofenolato mofetil
MPO	mieloperoxidase
MTX	metotrexate
mRNA	ácido ribonucléico mensageiro
<i>NF-κB</i>	nuclear factor kappa light chain of activated B cell
NK	células natural killer
Oligo dT	oligonucleotídeo com cauda de poli T (timina)
PCR	<i>polymerase chain reaction</i> ou reação em cadeia da polimerase
<i>Ripk2</i>	receptor-interacting serine-threonine kinase 2
RNA	ácido ribonucléico
<i>RNAse 3</i>	proteína catiônica eosinofílica

SVM-RFE *support vector machine – recursive feature elimination*

TBI total body irradiation

TCTH transplante de células tronco hematopoéticas

Th1 células T auxiliar tipo 1

Th2 células T auxiliar tipo 2

TMEM176B transmembrane protein 176B

TNF α fator de necrose tumoral alfa

TREC T cell receptor excise circle

TS iniciador *template switch*

Resumo

Introdução: Transplante alogênico de células tronco hematopoéticas (TCTH) é uma importante terapia para doenças hematológicas, mas o sucesso de uma porção de transplantes é limitado pela doença do enxerto versus hospedeiro (GVHD). Os fatores de risco conhecidos para GVHD agudo (aGVHD) não fornecem uma estimativa precisa do risco individual e não auxiliam na individualização da terapia. Até o momento, não existe método diagnóstico que permita prever aGVHD. A identificação de pacientes que desenvolverão aGVHD poderia permitir a individualização da terapia para uns e evitaria imunossupressão intensa para outros.

Objetivos: Revelar um classificador molecular preditivo de aGVHD. Descobrir genes diferencialmente expressos e analisar eventos precoces que desencadeiam aGVHD. Explorar categorias funcionais e tipos celulares relacionados ao desenvolvimento de aGVHD.

Casuística e métodos: Foram isolados, amplificados, marcados e co-hidridados com lâminas de microarray contendo 22.000 sondas, amostras de RNA mensageiro de 89 pacientes submetidos a TCTH HLA-idêntico mieloablativo ou de toxicidade reduzida, obtido de células mononucleares periféricas durante a enxertia medular. Os pacientes foram divididos em grupo treino e grupo teste, um modo utilizado para construir um modelo que discrimine pacientes com e sem aGVHD, e para testar o modelo em amostras independentes. Os genes informativos foram selecionados utilizando *recursive feature elimination*, seguido de sete diferentes algoritmos de classificação multivariados para estabelecer o classificador molecular no grupo treino. Os genes diferencialmente expressos entre amostras de pacientes com e sem GVHD foram submetidos a análise de enriquecimento das vias funcionais e agrupados de acordo com o perfil de expressão com células e tecidos através do SymAtlas.

Resultados: Encontramos um classificador molecular composto de 233 genes nas amostras do grupo treino, que foram selecionados baseados na mais precisa classificação. No grupo teste da amostra, cerca de 80% dos pacientes puderam ser

classificados. Para estes pacientes, o classificador mostrou uma acurácia preditiva de 75% (sensibilidade de 71% e especificidade de 78%). Analisando a anotação funcional dos genes diferencialmente expressos, observamos que em pacientes que desenvolveram aGVHD, houve aumento de expressão de genes da resposta antimicrobiana, transporte de gases, metabolismo de hemoglobina, além das alarminas. Nestas amostras também observamos a diminuição da expressão da *IL1* e outros genes envolvidos na via do *NF- κ B*. Vários genes super-expressos no período precoce do aGVHD foram associados a células precursoras.

Conclusões: Nossos resultados mostram que um classificador molecular é capaz de identificar pacientes sob alto risco de desenvolver aGVHD. Estabelecer métodos de diagnóstico preditivo para aGVHD é o primeiro passo para a individualização da estratégia terapêutica após TCTH. Além disso, os resultados da análise de enriquecimento funcional e a expressão de genes em diferentes populações celulares, sugerem que eventos precoces envolvendo múltiplas populações de células precursoras possam predefinir a interação futura entre o enxerto em desenvolvimento e o paciente.

Key-words: transplante alogênico de células tronco hematopoéticas, doença do enxerto contra o hospedeiro aguda, microarray, perfil de expressão gênica

1. Introdução

1.1. Descrição do problema: maior incidência de aGVHD de moderada a grave implica em maior morbi-mortalidade devido a lesão de órgão alvo, infecções e outras complicações relacionadas ao GVHD.

Transplante alogênico de células tronco hematopoéticas (TCTH) é modalidade de tratamento que promove eficácia terapêutica em pacientes com doenças hematopoéticas como leucemias agudas, leucemias crônicas, linfomas e depressões medulares. Para alguns pacientes, esta modalidade de tratamento permanece como única opção, oferecendo um potencial de cura ou controle a longo prazo da doença de base.

Estima-se que no ano de 2007 tenham sido realizados cerca de 1200 transplantes no Brasil, segundo dados do Ministério da Saúde, e em torno de 40.000 procedimentos pelo mundo⁽¹⁾, com um crescimento estimado de 10-15% ao ano. O número de transplantes tem aumentado significativamente e a fonte de células, tipo de doador, prevenção de doença do enxerto contra hospedeiro, novos regimes de condicionamento e melhora de terapias suportivas têm mudado nas última décadas⁽²⁾.

O sucesso de uma proporção destes transplantes é limitado pela doença do enxerto contra o hospedeiro (GVHD- *graft versus host disease*). Após o TCTH alogênico, complicações como toxicidade ao regime de condicionamento, infecções, e recaída da doença podem acontecer, mas a maior e mais importante complicação imunológica é GVHD. A forma aguda de GVHD (aGVHD) é complicação difícil de se manejar⁽³⁾ e ocorre em cerca de 30-60% dos pacientes submetidos a TCTH aparentando.

Esta condição é causada por linfócitos T do doador respondendo a aloantígenos expressos pelo receptor. Um dos fatores importantes para a emergência de aGVHD é a toxicidade induzida pelo regime de condicionamento, levando a lesão celular com exposição de neoantígenos, liberação de citocinas e aumento da apresentação antigênica às células T do doador. O reconhecimento de peptídeos alogênicos induz ativação, seguida de expansão, iniciando a resposta imunológica contra tecidos do hospedeiro⁽⁴⁾.

O regime de condicionamento tem papel importante na lesão tecidual, promovendo a liberação de citocinas pró-inflamatórias, tendo o tubo gastrointestinal como principal gatilho para a reação do enxerto contra o hospedeiro⁽⁵⁾. A redução da intensidade do regime de condicionamento foi inicialmente proposta com a intenção de reduzir a mortalidade relacionada ao transplante, permitindo a inclusão de pacientes mais idosos ou com co-morbidades. Apesar da nova estratégia, estudos não têm demonstrado menor mortalidade relacionada ao tratamento, menor incidência de GVHD agudo e crônico ou melhora da sobrevida^(6,7).

As características clínicas de aGVHD incluem lesões de pele, fígado e tubo gastrointestinal, graduados segundo Glucksberg⁽⁸⁾. Corticóides representam a terapia primária para aGVHD⁽⁹⁾, embora, ao tempo da reação imune clinicamente visível, 50% dos pacientes não responderão a corticoterapia, e apesar da introdução de terapias de segunda linha, em torno de 90% dos pacientes com GVHD agudo refratário morrerão dentro dos primeiros dois anos após transplante⁽¹⁰⁾. A incidência e gravidade de GVHD dependem de um número de fatores que incluem: tipo de profilaxia, tipo de doador, intensidade do regime de condicionamento, composição do enxerto e função tímica pré transplante⁽¹¹⁻¹⁶⁾.

A busca de preditores para eventos clínicos pós transplante é tema de estudo recorrente nos vários centros de transplante pelo mundo. A função tímica e sua correlação com eventos precoces e tardios após transplante é um destes objetos de estudo. Tem sido sugerido que a reconstituição e manutenção efetiva da imunidade conferida por células T é dependente da educação de precursores T no timo, como mostrado pela correlação entre aumento da produção tímica, aumento do número de células T imaturas, e amplo repertório de receptor de células T após transplante. Fatores que inibem função tímica podem incluir: idade, lesão tímica direta por quimioterapia ou radioterapia e GVHD⁽¹⁷⁻¹⁹⁾.

Desde a descrição por Douek et al dos círculos de excisão de receptor de células T (TRECs- TCR excision circles), a detecção de maior número de cópias de TRECs previamente ao transplante alogênico, denotando boa função tímica, tem sido

relacionada a menor incidência de GVHD agudo e crônico, menor incidência de infecções virais e bacterianas e maior sobrevida geral⁽¹⁶⁾. Nosso grupo estudou 62 pacientes, entre crianças e adultos, e também observamos o impacto dos baixos níveis de TRECs pré TCTH na diminuição da sobrevida geral e aumento da incidência de aGVHD grau II-IV, e de infecções bacterianas e virais, conforme demonstrado no artigo em anexo (Anexo 1). Nossos resultados corroboram a idéia de que a deficiência de função tímica pré transplante pode ser um preditor de risco para a ocorrência de aGVHD grau II-IV e de infecções.

Embora a presença de GVHD possa sugerir um efeito benéfico do enxerto contra o tumor, estudos observacionais identificaram GVHD como a principal causa de mortalidade não relacionada a recaída^(20,21). De acordo com dados do CIBMTR (Center for International Blood and marrow Transplant Research), GVHD representa 13% e 14% das causas de morte após transplante alogênico aparentado e não aparentado, respectivamente⁽²²⁾.

Os fatores de risco atualmente reconhecidos pela literatura⁽²³⁾ não são capazes de fornecer uma estimativa precisa do risco individual e dessa forma não permitem a individualização da terapia.

1.2. Uso de *microarrays* como ferramenta investigativa em hematologia e transplante de células tronco hematopoéticas

A conclusão do projeto Genoma, no final da década de 1990, possibilitou o surgimento de uma tecnologia que permite analisar a expressão diferencial de milhares de genes simultaneamente, os chamados *microarrays*^(24,25).

Esta tecnologia, apesar de relativamente nova, tem revolucionado as pesquisas em várias áreas da medicina, inclusive na hematologia, na medida em que amplia rapidamente a quantidade de informações disponíveis sobre a expressão gênica em diversas situações clínicas e experimentais.

A principal vantagem dos *microarrays*, descrito pela primeira vez em 1995 por Schena M, é o fato desse método não se basear no conhecimento prévio de quais genes estariam envolvidos no evento que se quer estudar, e por isso se ampliam as probabilidades de desvendar novos mecanismos envolvidos e possíveis biomarcadores para o diagnóstico do evento em estudo.

Esta técnica utiliza sequências de genes dispostas em arranjos (*array*), preparadas roboticamente e aplicadas em quantidades adequadas, seja de produto de PCR oriundos de clones de cDNA ou oligonucleotídeos gene-específicos sobre lâminas de vidro⁽²⁶⁾.

Tecnicamente, milhares de sondas são fixadas à superfície de uma lâmina de vidro e amostras de RNA (o alvo) são marcadas com fluoróforos para hibridação. Após a hibridação, um feixe de laser é utilizado para excitar o fluoróforo. A intensidade de hibridação é representada pela quantidade de fluorescência emitida, que fornece uma estimativa da quantidade relativa entre os diferentes transcritos que estão representados.

A plataforma de microarray mais utilizada no momento e que utilizamos neste estudo, são microarrays de oligos, em que sequências de 60-70 nucleotídeos específicos para genes previamente conhecidos, são sintetizadas e fixadas às lâminas. Os mRNA de amostras biológicas são submetidos a transcrição reversa e amplificação, se necessária, e depois marcados com fluoróforo (Cy3 ou Cy5). Após a hibridação, as fluorescências do Cy3 e Cy5 são mensuradas separadamente e capturadas em duas imagens, sendo então, sobrepostas e produzindo uma imagem composta. Esta imagem deve ser pré-processada, com adequada filtragem, normalização e subtração da fluorescência de fundo (*background*), antes dos valores de expressão serem analisados. Uma vez pré-processada, são utilizadas ferramentas adequadas para descoberta de classes, comparação de classes ou predição de classes, de acordo com o desenho pré estabelecido do estudo. Neste caso, classe é equivalente ao grupo de amostras ou de pacientes.

Em 1999, pela primeira vez, pesquisadores demonstraram que a análise da expressão diferencial de genes poderia ser usada para classificar um grupo de leucemias agudas⁽²⁷⁾. Baseado no padrão de expressão gênica de 55 genes diferencialmente expressos, leucemia mieloide aguda e leucemia linfóide aguda puderam ser distinguidas sem informações histológicas prévias. Este estudo demonstrou que classificações clinicamente úteis poderiam ser realizadas baseadas no perfil de expressão gênica.

Em pesquisa clínica, *microarrays* têm sido usados para prever sobrevida geral e resposta ao tratamento em diferentes doenças hematológicas⁽²⁸⁻³²⁾ e para diagnosticar rejeição aguda (um processo patológico que possui características similares ao GVHD agudo) em transplante de órgãos sólidos⁽³³⁻³⁵⁾.

Estudos de microarray com linfoma difuso de grandes células B por sua vez, identificaram 03 subgrupos moleculares distintos, aqueles com assinatura gênica de células B do centro germinativo (GCB), os de linfócitos B ativados (ABC) e o terceiro conhecido como tipo 3, que parecem ser derivados de diferentes estágios de diferenciação das células B, utilizando diferentes mecanismos oncogênicos, diferindo entre si em suas habilidades de serem curados por poliquimioterapia. Utilizando expressão diferencial de genes, pacientes que expressam assinatura molecular de células B de centro germinativo têm melhor sobrevida do que pacientes com assinatura de linfócitos B ativados^(28,29).

Um grande número de estudos tem abordado o perfil de expressão gênica e suas aplicações em classificar tumores, descobrir novas classes de doença, estimar prognóstico e prever resposta à terapia. Dentre estes tipos de análises, a predição de classes tem por objetivo construir um preditor ou classificador molecular para a ocorrência de eventos como recaída da doença de base, sobrevida geral e resposta à terapia⁽³⁶⁾.

Recentemente proposto por Guyon et al, o algoritmo chamado "Support Vector Machine Recursive Feature Elimination" (SVM-RFE) é um método estatístico que pode

ser utilizado para seleção de genes de determinado preditor⁽³⁷⁾. Este algoritmo ordena os genes de acordo com sua contribuição no desempenho do preditor. Ele utiliza uma combinação linear ponderada da expressão dos genes como um discriminador entre duas classes. Esta combinação linear é selecionada para maximizar a margem ou a distância entre a amostra pior classificada e o plano discriminador. A expressão de todos os genes é usada para o treino no algoritmo SVM, removendo genes que têm baixo peso na combinação linear e um novo classificador é desenvolvido utilizando os genes remanescentes. Após encontrar um novo discriminante linear, os genes que têm baixo peso para o novo discriminante são removidos e um novo classificador é criado. O processo continua iterativamente até que o número de genes previamente selecionado seja atingido. Este método faz parte de um sub-campo da inteligência artificial chamada aprendizagem de máquina, dedicado ao desenvolvimento de algoritmos e técnicas que permitam ao computador aprender, isto é, que permitam ao computador aperfeiçoar seu desempenho em alguma tarefa.

O conceito de sobre-ajuste ou *overfitting* em aprendizado de máquinas é importante pois, durante a criação de um preditor, a máquina pode ajustar aleatoriamente as características do grupo treino, não tendo relação causal com o evento buscado. Neste processo de sobre-ajuste, o desempenho em amostras treinos tende a melhorar enquanto que em amostras teste independentes, os resultados tendem a piorar. Desta forma, neste tipo de análise, é essencial que os dados usados para avaliar a acurácia preditiva do classificador sejam diferentes dos dados que geraram a lista de genes que constitui o preditor.

Desta forma, dois grupos de amostras devem ser criados, um chamado de grupo treino ou *training set*, no qual é criado e selecionado o preditor, e outro chamado de grupo teste ou *test set*, grupo este em que o preditor é aplicado e avaliado⁽³⁸⁾.

1.3. Predição em GVHD agudo

Em 2007, um estudo utilizando a técnica de *microarray* foi publicado, utilizando o perfil de expressão do doador para prever a ocorrência de GVHD após transplante HLA-idêntico de células tronco periféricas ou medula óssea⁽³⁹⁾. O investigador analisou o perfil de expressão gênica em linfócitos T CD4+ e CD8+ provenientes de sangue periférico de doadores. Nesta análise, foi encontrado um conjunto de genes cujo classificador molecular foi capaz de prever GVHD (agudo e crônico) com acurácia de 65%, e combinando pares de genes, foram capazes de prever o desenvolvimento de GVHD crônico com um acurácia de 77%. Desta forma, a acurácia atribuída a predição de aGVHD deve-se mais à capacidade de prever GVHD crônico, não existindo, portanto, um classificador para aGVHD. Embora estes resultados sejam promissores (este é o primeiro e único estudo que explorou a predição de GVHD utilizando *microarray* em doadores), existem limitações. Primeiro, o autor não mostra um conjunto de genes preditivo para aGVHD sozinho, talvez porque o número limitado de doadores/pacientes (N:50) não tenha sido capaz de gerar informações suficientes que permitissem o estabelecimento de um classificador robusto e confiável, ou talvez porque analisaram somente células T CD4 e CD8, e evidências recentes apontam que células B também estão envolvidas em GVHD, e o perfil de expressão dessas células pode ser bastante informativo⁽⁴⁰⁾. Segundo, a análise de predição foi realizada em amostras colhidas de doadores de células tronco hematopoéticas periféricas ou medula óssea antes do transplante. Ainda que o perfil de expressão das células T do doador reflita seu status de ativação imune, isto pode não estar relacionado à capacidade do doador em montar uma resposta do enxerto contra o hospedeiro após o contato com aloantígenos.

Nosso grupo recentemente utilizou com sucesso a tecnologia de *microarray* para prever reação do hospedeiro contra o enxerto em pacientes submetidos a transplante cardíaco. Neste trabalho, conseguiu-se determinar uma assinatura molecular que permitiu a detecção precoce de episódios de rejeição antes do diagnóstico histológico⁽³⁵⁾.

Como aGVHD é uma reação imune similar (células T respondendo a aloantígenos), nós postulamos que deveria ser possível encontrar uma assinatura molecular precocemente após o transplante de medula óssea, que permitiria prever a ocorrência de aGVHD, antes dos sintomas clínicos aparecerem. A identificação de uma assinatura gênica predizendo aGVHD abriria caminho para estudos com intensificação de profilaxias para alguns, evitaria imunossupressão excessiva para outros que não estivessem sob risco, além da possibilidade de identificar pacientes potencialmente refratários à corticoterapia. Diante destas justificativas, por não existir método capaz de prever quais pacientes sofrerão aGVHD, e ainda porque fatores de risco reconhecidos pela literatura não permitem estimar o risco individual, nós desenhamos este estudo. Dessa forma, no presente estudo, foi avaliado o perfil de expressão gênica com o emprego de microarrays em 89 pacientes submetidos a transplante alogênico HLA-idêntico, com o objetivo de descobrir uma assinatura molecular capaz de prever a ocorrência de aGVHD. Além disso, através da análise de anotação funcional dos genes diferencialmente expressos, exploramos as vias biológicas envolvidas na fase precoce de aGVHD.

2. Objetivos

1. Revelar um classificador molecular preditivo de doença do enxerto contra o hospedeiro aguda durante a enxertia medular.
2. Descobrir genes diferencialmente expressos e analisar os eventos precoces que desencadeiam aGVHD.
3. Explorar categorias funcionais e tipos celulares relacionados ao desenvolvimento de aGVHD.

3. Casuística e métodos

3.1. Casuística

Foram recrutados prospectivamente e consecutivamente 103 pacientes submetidos a transplante alogênico HLA-idêntico aparentado para o presente estudo. Destes, 89 pacientes participaram do estudo e 14 foram excluídos (critérios de exclusão abaixo). Do total de pacientes recrutados, 53 foram acompanhados na Disciplina de Hematologia e Hemoterapia do Departamento de Medicina da Universidade Federal de São Paulo (UNIFESP), 50 no Setor de Hematologia do Hospital Santa Marcelina, seguindo-se os mesmos critérios de inclusão e de exclusão estabelecidos, após cuidadosa avaliação clínica de cada paciente. Foram recrutados pacientes submetidos a transplante entre o período de maio de 2004 a agosto de 2007.

Foram incluídos pacientes de ambos os sexos, com idade entre 15 e 70 anos e com intervalo maior ou igual a 7 dias entre a enxertia medular e o início da doença do enxerto contra o hospedeiro aguda. Pacientes que tiveram aGVHD foram incluídos independente do tempo de seguimento. Pacientes com quadro de sepse ou doença veno-oclusiva manifesta durante a recuperação medular foram excluídos. Na análise de predição da doença do enxerto contra hospedeiro aguda, excluímos pacientes que não tiveram aGVHD e que foram a óbito por infecção, doença veno-oclusiva hepática ou outras complicações nos primeiros 100 dias, além dos pacientes que não foram capazes de entender o termo de consentimento livre e esclarecido.

Todos os pacientes foram acompanhados por, pelo menos, 100 dias após transplante, com avaliação ambulatorial 3 vezes por semana nos primeiros 100 dias. Para estudo de expressão gênica foi utilizada uma amostra por paciente.

Gradação de doença do enxerto foi avaliada por dois médicos independentes e casos duvidosos foram consenso com um terceiro médico. Foi considerado doença do enxerto Grau I ou de órgão único, se confirmado por biópsia. O critério utilizado para classificação de aGVHD foi o critério clínico de Glucksberg modificado⁽⁴⁾, sendo documentada a graduação máxima atingida pelo aGVHD. O trabalho foi aprovado

pelos Comitês de Ética da Universidade Federal de São Paulo e do Hospital Santa Marcelina (Anexo 7). Todos os pacientes assinaram o Termo de Consentimento Livre e Esclarecido (Anexo 8).

3.2. Técnicas de transplante utilizadas

Todos os transplantes realizados foram HLA-idênticos de doadores irmãos com compatibilidade HLA nos seguintes locus : HLA classe I (HLA-A e HLA-B) realizados por métodos sorológicos e HLA classe II (DR e DQ) realizados por métodos moleculares de alta resolução.

Foram considerados transplantes mieloablativos aqueles com os seguintes regimes de condicionamento: Bussulfano (16mg/Kg) associado a ciclofosfamida(120mg/Kg), Irradiação corporal total (1200 cGy) associado a ciclofosfamida(120mg/kg), bussulfano (16mg/Kg) associado a fludarabina(120mg/m²) e carmustina (300 mg/m²), citarabina (1600 mg/m²), etoposide (800 mg/m²) associado a melfalan (140 mg/m²). A profilaxia de GVHD foi realizada com ciclosporina 3mg/Kg endovenosa, iniciado no D-1 e retirada progressiva após D+60 e suspensão no D+180. Metotrexate foi utilizado nos dias +1 (15 mg/m²), +3, +6 e +11 (10mg/m²)⁽⁴¹⁾.

Foram considerados transplante de toxicidade reduzida os seguintes regimes de condicionamento^(42,43): melfalan (140 mg/m²), fludarabina(120 mg/m²) associada a irradiação corporal total (200cGy), ciclofosfamida (120mg/Kg) associada a melfalan (120 mg/m²), fludarabina (120 mg/m²) associada a bussulfano (8mg/kg) e linfoglobulina antitimocítica (40mg/Kg). Para o regime de condicionamento com fludarabina e irradiação corporal total foi utilizada a profilaxia com ciclosporina 12,5 mg/Kg via oral iniciando no D-3 do transplante e micofenolato mofetil 15mg/Kg no D0, com retirada do micofenolato no D+28 e ciclosporina entre D+60 e 90, dependendo do quimerismo⁽⁴⁴⁾. Para os outros regimes, a profilaxia foi igual ao regime mieloabaltivo. A concentração sérica da ciclosporina foi avaliada semanalmente e mantida entre 150 – 300 ng/ml em ambos tipos de transplantes.

Profilaxia antibiótica foi realizada igualmente entre os dois regimes: sulfametoxazol associado a trimetropim para profilaxia de *Pneumocystis jiroveci*, mantido até o término da imunossupressão ou até seis meses após o controle da doença do enxerto contra hospedeiro crônica. Aciclovir (10mg/Kg 3x/dia) e fluconazol (200mg/dia) foram iniciados durante o regime de condicionamento na forma endovenosa, depois da alta foram modificados para aciclovir 800mg/dia e fluconazol 200mg/dia via oral e então suspensos nos primeiros 100 dias, exceto se a doença do enxerto contra o hospedeiro necessitasse de prolongada imunossupressão⁽⁴⁵⁾.

Antigenemia para citomegalovírus(CMV) foi realizada semanalmente até o D+100. Após os primeiros 100 dias, a monitorização foi realizada baseada no risco de cada paciente. Foi considerada reativação de CMV se positividade presente ≥ 2 células/150.000 leucócitos, sendo instituído terapia com ganciclovir. GVHD crônico foi classificado de acordo com critérios de Seattle⁽⁴⁶⁾.

As doenças foram categorizadas antes do transplante da seguinte forma: grupo de doenças não malignas: anemia aplástica; grupo de baixo risco: síndrome mielodisplásica com anemia refratária, síndrome mielodisplásica hipocelular, leucemia mielóide crônica em fase crônica, leucemias aguda em primeira remissão, linfomas em remissão; grupo intermediário: leucemias em segunda remissão, leucemia crônica em fase acelerada ou 2ª fase crônica, hemoglobinúria paroxística noturna, doença mieloproliferativa inclassificável; grupo alto risco: leucemia mielóide crônica em crise blástica, síndrome mielodisplásica com excesso de blastos, leucemia mielomonocítica, leucemia secundária a mielodisplasia, leucemia bifenotípica, leucemias agudas em atividade, linfomas ou mieloma múltiplo em atividade⁽⁴⁷⁾.

3.3. Coleta de amostras e processamento

3.3.1 Processamento do material

As células mononucleares de sangue periférico foram separadas por Ficoll-Hypaque®, conforme orientações do fabricante, a partir de 10 ml de sangue colhido em heparina durante a enxertia medular após transplante. Foi considerado enxertia medular contagens absolutas de neutrófilos $> 500/\text{mm}^3$ por três dias consecutivos, sendo considerado o primeiro dia como a data da recuperação medular e plaquetas $> 20.000/\text{mm}^3$ sustentada por, pelo menos, 7 dias consecutivos sem transfusão⁽⁴⁸⁾.

3.3.2. Extração de RNA total

Os RNAs foram extraídos imediatamente após separação de células no Ficoll, utilizando Trizol (Invitrogen®), segundo informações do fabricante e armazenados a -80°C até o momento do uso.

3.3.3. Avaliação da Qualidade e Quantidade do RNA total

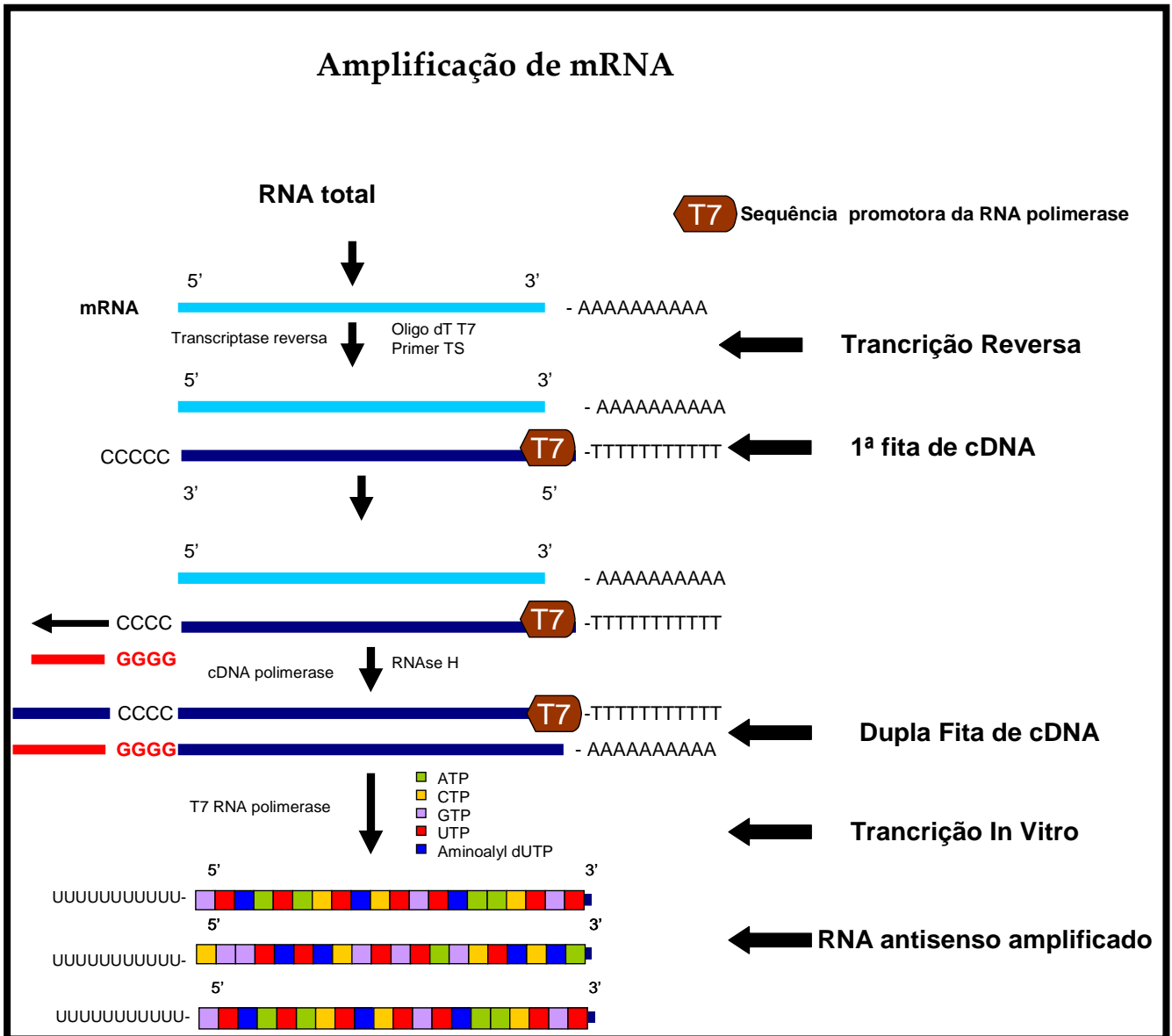
A quantidade de RNA total foi analisada por espectrofotômetro de leitura ultra violeta (GeneQuant II- Pharmacia Biotech®), com leitura realizada na faixa de 280 e 260 nm. Foram consideradas amostras para inclusão no estudo, aquelas com absorbância entre 0,1 e 1,0 e ratio (280 e 260 nm) entre 1,6 -2,0. A qualidade do RNA foi analisada por equipamento Bioanalyzer Agilent 2100 através de separação eletroforética utilizando kit LabChip RNA6000 (Agilent®). Foram consideradas amostras de boa qualidade aquelas que mantinham relação 28S/18S entre 1,6-2,0.

3.3.4 Amplificação do RNA mensageiro (mRNA)

Para os experimentos de *microarray*, os RNAs mensageiros foram amplificados em uma etapa (Figura 1). A partir de 1 μg de RNA total, na presença de oligo dt(24) T7 (5'-AAA CGA CGG CCA GTG AAT TGT AAT ACG ACT CAC TAT AGG CGC TTT TTT TTT TTT TTT TTT TTT TTT-3) e de um iniciador ``template switcher`` (5'- AAG

CAG TGG TAA CAA CGC AGA GTA CGC GGG-3'), as amostras foram submetidas a transcrição reversa. A utilização do oligo dT(24)T7 permite a introdução da região promotora do gene da RNA polimerase no cDNA obtido após a transcrição reversa do RNA total. O iniciador *template swichter* (TS) permite a extensão da cadeia molde até o fim, através da adição de uma cauda poli G. A segunda fita de cDNA foi sintetizada na presença de cDNA polimerase e seguida de precipitação. A dupla fita de cDNA foi submetida a transcrição *invitro* na presença de RNA polimerase Phago T7. Além dos nucleotídeos necessários, foi adicionado um nucleotídeo modificado (aminoalil_dUTP) que permite a marcação do RNA amplificado (aRNA) com os corantes Cy3 ou Cy5 numa etapa posterior^(35,49,50). Para amplificação do mRNA, foi utilizado somente um *round* de reação.

Figura 1- Ilustração do método utilizado para amplificação de mRNA



3.4. Método de *Microarray*

As lâminas de *microarray* utilizadas neste projeto foram preparadas por NIAID *Microarray Facility*, NIH, Estados Unidos. Estas lâminas contêm impressos oligonucleotídeos de 70 mers correspondentes a 22.000 genes humanos sintetizados por Qiagen/Operon (Estados Unidos).

Antes da realização da hibridação, a quantidade de RNA antisense amplificado foi analisado por espectrofotômetro de leitura ultra violeta (GeneQuant II- Pharmacia Biotech) e a qualidade do RNA foi analisada por equipamento Bioanalyzer Agilent 2100 através separação eletroforética utilizando kit LabChip RNA6000 (Agilent).

Para cada experimento de hibridação, foi utilizada uma amostra-referência comum. Esta amostra-referência foi obtida a partir de um *pool* de RNA de todas as amostras utilizadas nos experimentos. Os aRNAs marcados com Cy3 (amostra) e Cy5 (amostra-referência) são misturados e utilizados conjuntamente na hibridação a uma lâmina de *microarray*. Após lavagem, as lâminas foram escaneadas utilizando o scanner de *microarray* GenePix 4000B (Axon Instruments). Os valores de expressão gênica foram obtidos pelo software GenePix Pro 6.0 (Axon Instruments). Estes foram importados para softwares de análise de dados de expressão obtidos por microarrays (Figura 2).

3.4.1. Escaneamento das lâminas de *Microarray*:

Utilizando o aparelho Scanner Axon 4000B (Axon Instruments -Molecular Devices®), as lâminas foram escaneadas com leitura realizada com resolução de 10µm, cujas imagens podem ser vistas na Figura(3A e 3B). Este equipamento captura intensidade luminosa de comprimentos de onda, 532nm para o canal verde e 633nm para o canal vermelho em cada *spot*. A leitura gerou dois arquivos com imagens separadas para cada canal (Cy3 e Cy5) e uma imagem com os dois canais sobrepostos.

Figura 2. Representação esquemática de um experimento de *microarray*

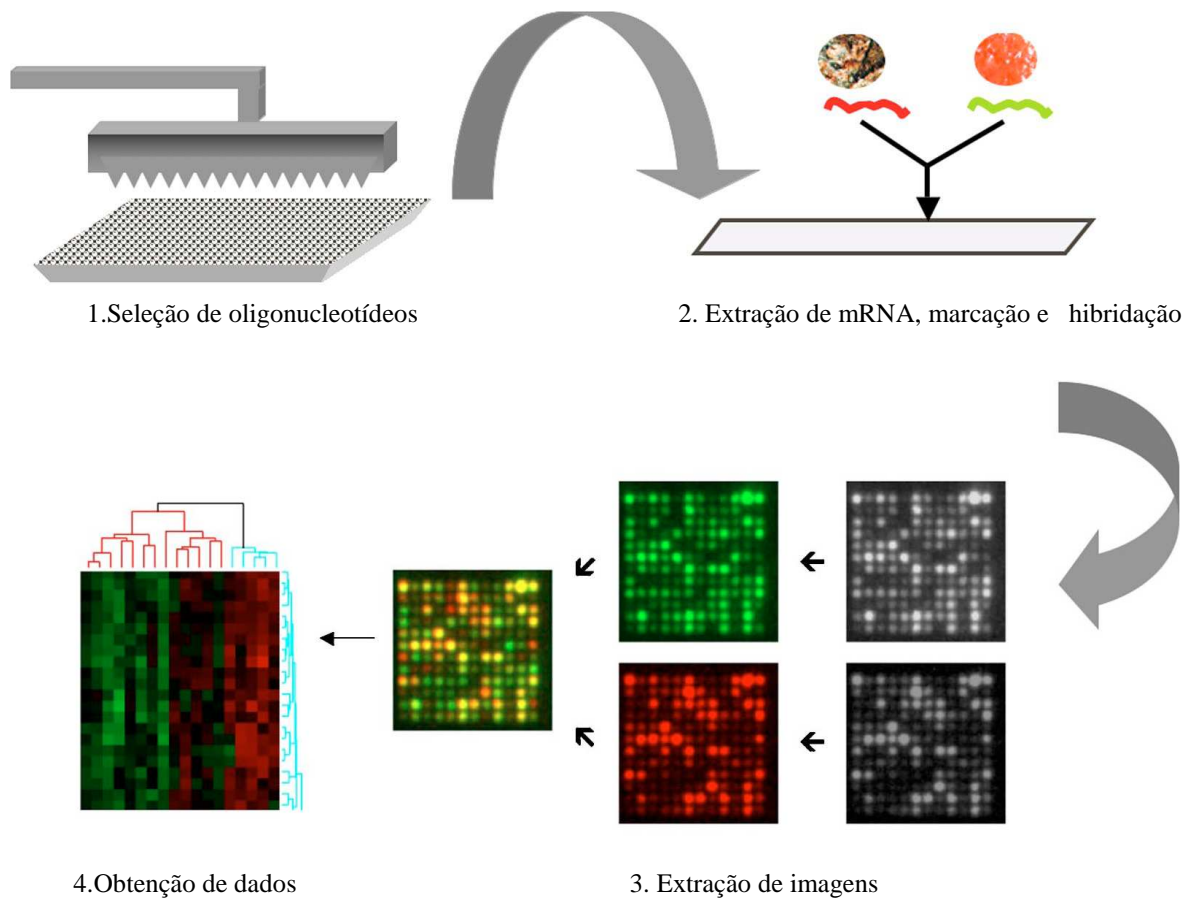
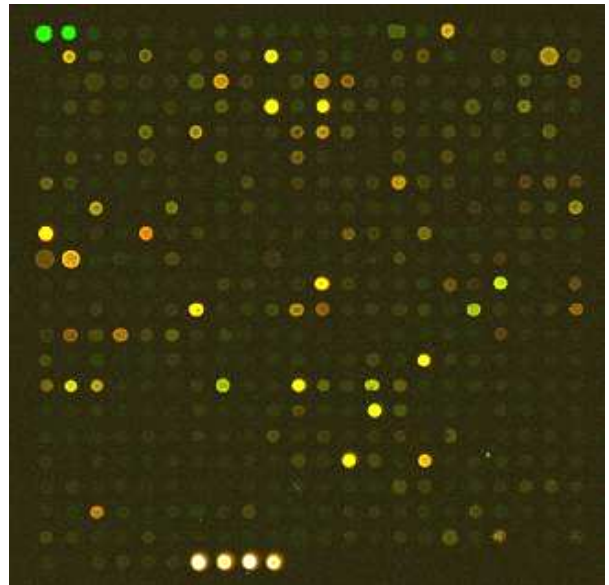


Figura 3. **A)** Imagem de uma lâmina de *microarray* após escaneamento com 10 μ m em Scanner Axon 4000B (Axon Instruments -Molecular Devices®). **B)** Imagem aumentada de um bloco de aproximadamente 400 genes.

A)



B)



3.4.2. Filtragem dos dados gerados

As imagens obtidas pelo Scanner foram transformadas em dados numéricos por meio do programa GenePix Pro 6.0. Este programa converte as imagens capturadas pelo scanner nos dois canais (verde e vermelho). A intensidade dos *spots* é medida e uma tabela com a intensidade e ruído de fundo (*background*) foi gerada. Uma vez adquiridos, os arquivos foram submetidos ao mAdb μ array Center Home-Nacional Cancer Institute/ National Institute of Health (<http://nciarray.nci.nih.gov/>). A análise foi realizada, utilizando o software BRB Array Tools (versão 3.7.0) desenvolvido pelos Drs. Richard Simon e Amy Peng Lam (<http://linus.nci.nih.gov/BRB-ArrayTools.html>).

Os seguintes critérios de filtragem foram utilizados: diâmetro de *spots* entre 10 e 300 μ m, valores de intensidade menor que 50 (canal verde) e menor que 20 (canal vermelho) em ambos canais após subtração do “background”, *spots* sem sinal e genes encontrados, em pelo menos, 50% das lâminas de array. Os valores mínimos de intensidade de cada canal foram determinados baseando-se na expressão de 190 controles negativos contendo sequências randômicas não humanas.

3.4.3. Normalização da expressão gênica

Durante a execução da técnica de *microarray*, a normalização é necessária para retirar variações sistemáticas que ocorrem durante a execução da técnica. O ajuste das intensidades de hibridações permitiu que a análise comparativa entre lâminas pudesse ser feita, garantindo melhor significado biológico dos resultados. Isso se deve a algumas razões próprias do experimento que incluem: quantidades diferentes de RNA entre amostras, variação na incorporação dos corantes, variação na quantidade de ruído de fundo (*background*) gerado pela deposição de corantes na lâmina de array⁽⁵¹⁻⁵³⁾.

Como a amostra do paciente foi marcada com Cy3 (que é representada posteriormente pela luz verde artificial) e a amostra referência, representada pelo pool de RNAs dos pacientes, marcada com Cy5 (representada posteriormente pela luz vermelha), o valor de expressão de determinado gene em relação a referência foi

calculado da seguinte forma: calculamos a razão G/R, sendo G o valor de intensidade do canal verde e R o valor de intensidade do canal vermelho para um dado gene, seguido de transformação logaritma de base 2 desta razão. Desta forma, o valor de expressão (T) de determinado gene (i) em dada amostra, em relação à referência, foi dada por:

$$T_i = \text{Log}_2 G_i/R_i$$

O Log_2 da razão das intensidade do Cy3 e Cy5, gerou um espectro contínuo e linear de valores. Desta forma, se o valor de expressão T_i de um gene i na amostra n for igual ao valor de expressão T_i da amostra referência, teremos uma razão de 1, com o respectivo $\text{Log}_2(1)=0$; se o gene for duas vezes mais expresso na amostra, a razão será 2 e $\text{Log}_2(2)=1$; e para duas vezes mais expresso na referência, a razão será $1/2$ e $\text{Log}_2(1/2)=-1$ ⁽⁵¹⁾.

Para normalização das intensidades obtidas, utilizamos o método de normalização Lowess (*Locally weighted linear regression*). Este método minimiza o efeito decorrente da intensidade, calculando um fator de ajuste por regressão linear ponderada local. Desta forma, calcula-se os dados centralizados, representados por $\log_2(R/G)$ menos o fator de ajuste. Estes dados centralizados deverão estar em torno do zero do eixo das ordenadas [$\log_2(R/G)$]. Desta forma, os dados foram normalizados em relação à mediana do logaritmo da razão das intensidades de ambos os canais (verde e vermelho) de expressão gênica de cada lâmina, tornando as lâminas comparáveis entre si (Figura 4).

O gráfico do tipo *MA plot* é utilizado para visualização dos dados brutos de *microarray*, sendo M (eixo y) o logaritmo da razão da intensidade ($\text{Log}_2(R_i/G_i)$) e A (eixo x) a média dos logaritmos de intensidade [$(\text{Log } R_i + \text{Log } G_i)/2$]. Os gráficos A demonstram dados de expressão obtidos de três lâminas antes da normalização e os gráficos B, da figura 5 demonstram dados obtidos das mesmas três lâminas após normalização com método Lowess, já com as devidas correções de variações sistemáticas em função das intensidades. Como se pode ver no exemplo ilustrado, o gráfico *MA* permite verificar que a dispersão aleatória da razão R/G pode variar

conforme a intensidade de emissão de sinal do *spot*. Após a normalização dos dados, este viés é corrigido(Figuras 4 e 5)

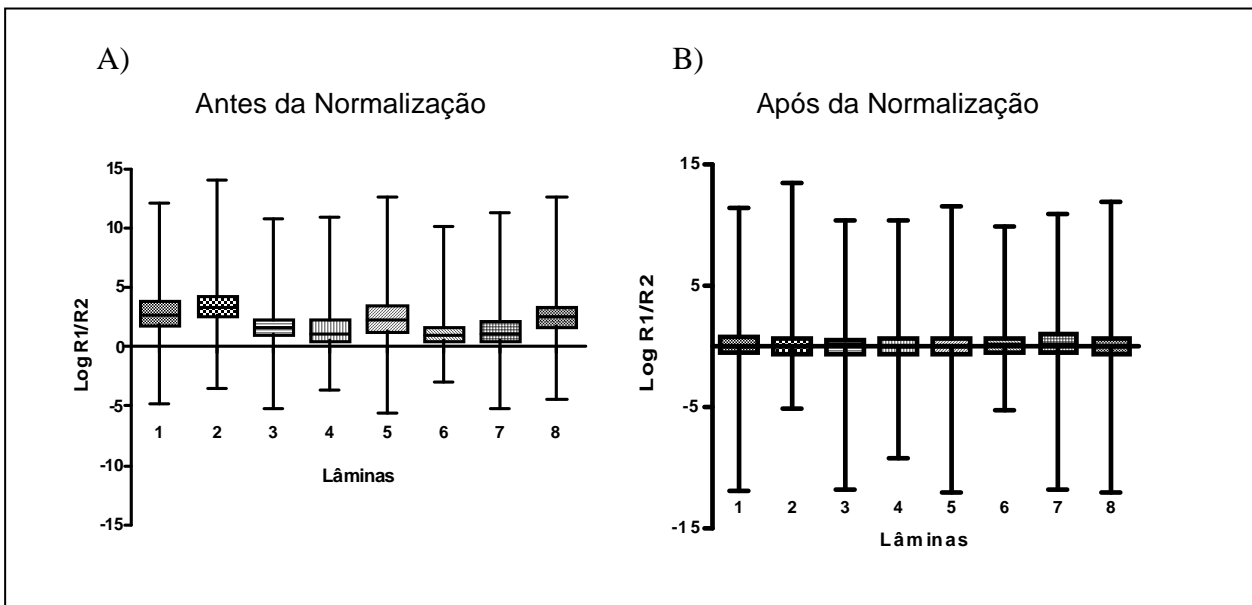


Figura 4. Gráficos do tipo *box-plot* mostrando o Logaritmo da razão entre as intensidades do canal verde (Cy3) e canal vermelho (Cy35) de oito lâminas de *microarray* antes e após o procedimento de normalização. Após a normalização, as medianas do sinal de cada lâmina estão mais próximas, tornando-as comparáveis entre si.

Figura 5: A) Gráfico do tipo *MA plot* de três lâminas de microarray antes da normalização.

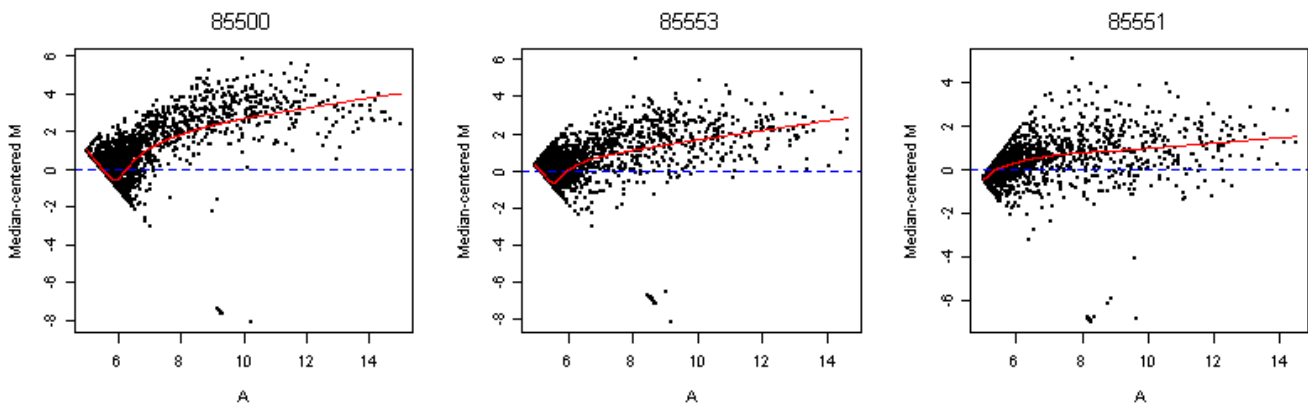
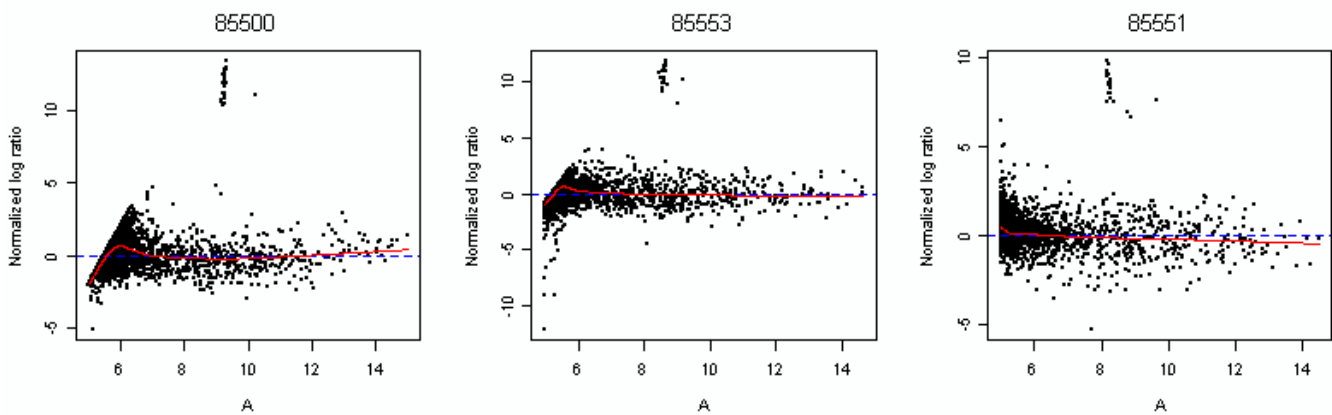


Figura 5:B) Gráfico do tipo *MA plot* de três lâminas de *microarray* após a normalização.



3.5 Métodos estatísticos

3.5.1. Métodos estatísticos para análise de dados clínicos

Para a análise de sobrevida, utilizamos o método de Kaplan-Meier e teste de Log-rank. O modelo Proporcional de Cox foi utilizado para identificar fatores de risco independentes para a sobrevida geral dos pacientes transplantados, revisados por Klein et cols⁽⁵⁴⁾. As variáveis incluídas no modelo foram aquelas cuja significância foi $\leq 0,10$ na análise univariada. Para estas análises foi utilizado o Software SPSS 10.0.

Os dados demográficos e as características clínicas nos dois grupos (com aGVHD e sem aGVHD) foram comparados utilizando teste t de Student ou Teste de Mann-Whitney para variáveis contínuas. O teste exato de Fisher foi utilizado para avaliar dados categóricos.

3.5.2. Método estatístico para geração e validação do classificador molecular

Análises dos dados brutos foram realizados utilizando o BRB Array Tools desenvolvido por Richard Simon e Amy Peng Lam (<http://linus.nci.nih.gov/BRB-ArrayTools.html>). Para classificar as amostras baseados no perfil de expressão, utilizamos ferramentas de predição de classes, que incluíram sete métodos multivariados de classificação com validação cruzada dos resultados e o uso de *recursive features elimination*. Os métodos foram: *compound covariate predictor*, *diagonal linear discriminant analysis*, *1- and 3-nearest neighbor predictor*, *nearest centroid predictor*, *support vector machine* e *Bayesian compound covariate predictor*.

Para prevenir um classificador com sobre-ajuste ou *over-fitting*, nós estimamos o erro do preditor para cada modelo usado, utilizando a validação cruzada do tipo *Leave-one -out cross-validation* como descrito por Simon e cols⁽⁵⁵⁾. Para cada cenário de treino (*training set*), o modelo inteiro foi construído e repetido, incluindo o processo de seleção dos genes.

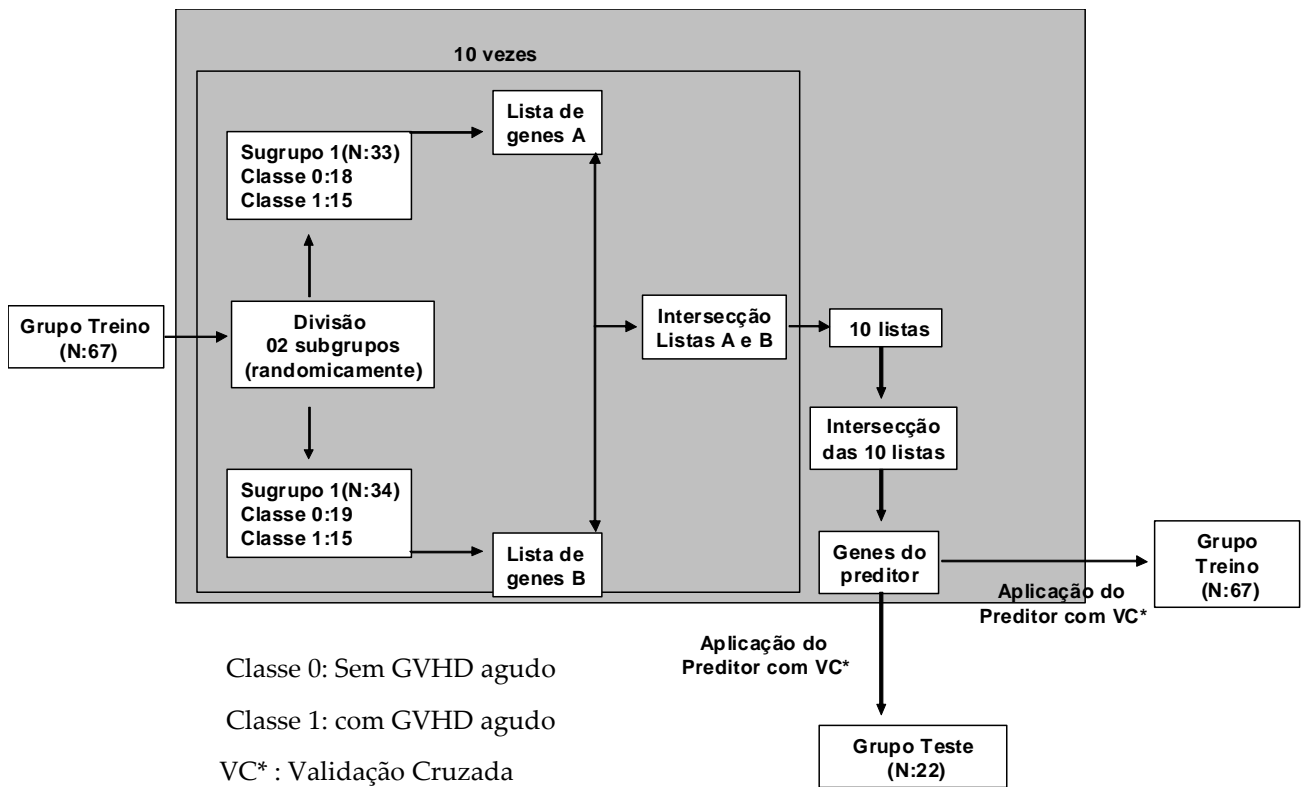
Nós dividimos o grupo de 89 pacientes em: grupo treino (N=67) e grupo teste (N=22). Para a geração do preditor nas 67 amostras (figura 6), dividimos esta amostra

em dois subgrupos de treino: subgrupo 1 (N=33 amostras); 18 sem aGVHD e 15 com aGVHD, subgrupo 2 (N=34 amostras); 19 sem aGVHD e 15 com aGVHD, e esta divisão foi realizada 10 vezes, mas preservando a proporção de pacientes. Para cada subgrupo de treino, utilizamos *recursive features elimination*, selecionando 1000 genes para a construção do preditor. Fizemos a intersecção destas duas listas, repetindo dez vezes e gerando dez listas diferentes. Estas dez listas foram submetidas a intersecção novamente e uma lista definitiva de genes foi obtida compondo o preditor.

Nós também avaliamos se a taxa de erro da validação cruzada para qualquer modelo foi significativamente menor do que a esperada por uma predição randômica. As classes (com aGVHD e sem aGVHD) foram randomicamente permutadas 1000 vezes e o processo de predição repetido.

Entre os sete métodos estatísticos, nós consideramos apenas os resultados que produziam uma permutação com valor de $p < 0,001$. Para a classificação de amostras, consideramos classificáveis pelo nosso método, aquelas que possuíam concordância em, pelo menos, seis métodos. Amostras com concordância em 5 métodos ou menos não foram classificáveis pelo nosso algoritmo, portanto, nesta situação, nada pode-se dizer sobre a classificação da amostra.

Figura 6. Representação esquemática da seleção dos genes para preditor e teste do preditor.



3.5.3. Análise de genes diferencialmente expressos

Para análise de genes diferencialmente expressos que compuseram o preditor, utilizamos a ferramenta de comparação de classes (*class comparison tool*), que permite a seleção dos genes diferencialmente expressos e retorna a probabilidade do erro α para cada gene, e também a probabilidade de encontrar por acaso o número de genes selecionados como diferencialmente expressos entre as classes no teste t univariado. Para a seleção de genes, utilizamos o limite de 10% para a taxa de descobertas falsas (FDR, ou *false discovery rate*)⁽⁵⁶⁾.

Para o agrupamento dos genes diferencialmente expressos em relação aos tecidos e linhagens celulares humanos já publicados na literatura⁽⁵⁷⁾, utilizamos os dados do SymAtlas disponível no site <http://wombat.gnf.org/SymAtlas/>. O método de clusterização utilizado foi a distância Euclidiana e *average linkage* que é uma análise implementada no software MeV (<http://www.tm4.org/mev.html>).⁽⁵⁸⁾

3.6. Análise funcional dos genes: a busca em banco de dados

Para a anotação funcional dos genes diferencialmente expressos, utilizamos um algoritmo disponível no programa ExPlain (BIOBASE, Germany), disponível no site <http://www.biobase-international.com/pages/index.php>. O algoritmo testa o enriquecimento de categorias do *Gene Ontology* dentro de nossa lista⁽⁵⁹⁾. Esta análise é realizada para cada grupo funcional separadamente. A significância escolhida foi $p < 0,01$ bicaudal. Para correção de teste de múltiplas hipóteses, o algoritmo utiliza o método de Benjamini/Hochberg⁽⁵⁶⁾.

Para verificar a existência de conexões, ou redes de comunicação bioquímica, entre genes sub-expressos ou super-expressos nas amostras de pacientes que evoluíram para aGVHD, nós utilizamos o software Ingenuity Pathway Analysis, disponível no site : [http://: www.ingenuity.com](http://www.ingenuity.com).

4. Resultados

4.1. Características dos pacientes

- Pacientes da UNIFESP:

No período de 05/2004 a 08/2007 foram encaminhados ao Setor de Transplante de Medula Óssea da Disciplina de Hematologia da UNIFESP, 53 pacientes .

Do total de 53 pacientes recrutados, 07 (13,2%) foram excluídos pela presença dos seguintes critérios de exclusão: 02 não aceitaram participar do estudo, 01 paciente com leucemia mielóide aguda refratária recuperou após transplante com doença em atividade, 01 paciente com linfoma do manto faleceu de sepse por *Pseudomonas aeruginosa* no D+12 após transplante, 01 paciente faleceu de meningite bacteriana no D+24 , 01 paciente faleceu de pneumonite por citomegalovírus no D+66, 01 paciente teve enxertia medular com emergência de GVHD hiperagudo no D+14. Foram então incluídos 46 pacientes (86,7% dos casos) da UNIFESP.

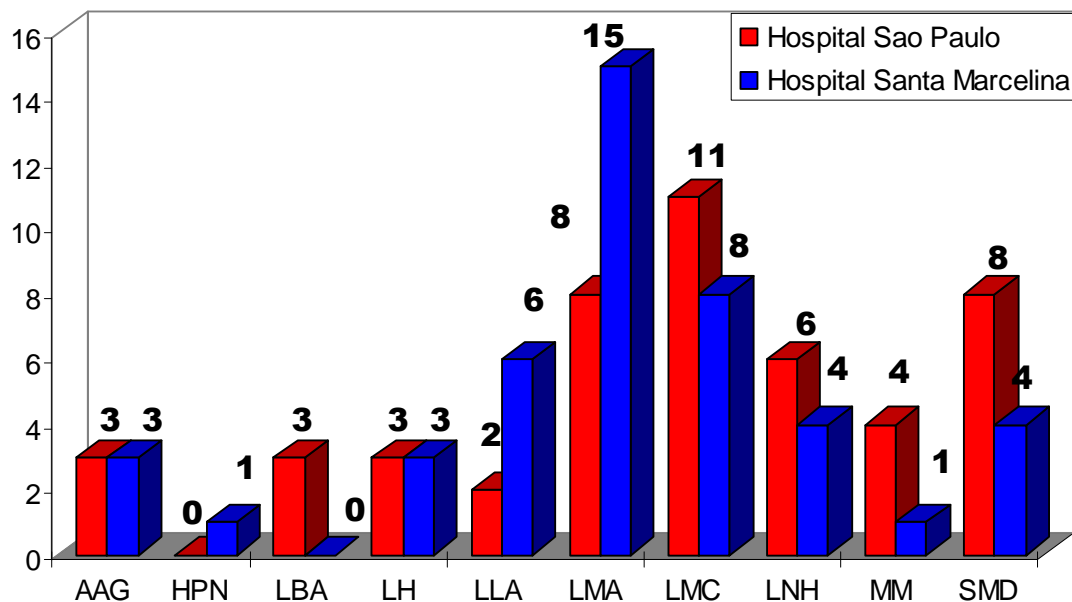
- Pacientes de outra instituição: Hospital Santa Marcelina:

Adicionalmente aos pacientes da UNIFESP, foram ainda recrutados 50 pacientes do Serviço de Transplante de Medula Óssea do Hospital Santa Marcelina, que obedeceram aos mesmos critérios de seleção.

Do total de 50 pacientes recrutados entre 05/2004 a 08/2007, 43 (86%) pacientes foram incluídos no estudo. Foram excluídos 07 (14%) pela presença dos seguintes critérios de exclusão: 01 não aceitou participar do estudo, 01 paciente com leucemia mielóide aguda faleceu de doença veno-oclusiva hepática no D+14 após transplante, 01 paciente com Linfoma de Hodgkin recidivado após transplante autólogo de medula óssea foi submetido a transplante alogênico e faleceu de sepse no D+9 após transplante, 01 paciente com anemia aplástica grave faleceu de sepse secundária a foco pneumônico no D+12 após transplante, 01 paciente faleceu no D+68 por aspergilose pulmonar, 01

paciente faleceu de choque séptico no D+52, 01 paciente teve enxertia medular e emergência de GV HD hiperagudo.

Figura 7: Distribuição dos 89 pacientes incluídos de acordo com o Hospital e o tipo de doença.



AAG -Anemia Aplástica Grave; HPN - Hemoglobinúria Paroxística Noturna; LBA- Leucemia Bifenotípica; LH -Linfoma de Hodgkin ; LLA -Leucemia Linfóide aguda; LMA -Leucemia Mielóide Aguda; LMC-Leucemia Mielóide Crônica; LNH - Linfoma Não Hodgkin; MM-Mieloma Múltiplo; SMD - Síndrome Mielodisplásica

4.2. Impacto da ocorrência de aGVHD e de fatores que interferiram na sobrevida geral ao final do primeiro ano

Neste estudo envolvendo 89 pacientes de dois centros de transplante, com um seguimento mediano de 13,8 meses (1,2 - 42,6), 40 (44,8%) pacientes desenvolveram aGVHD (grau I-IV). A emergência de aGVHD ocorreu com uma mediana de 33 dias (19 - 88 dias) após o transplante, com a seguinte frequência baseada na Escala de Gluksberg: 04 pacientes tiveram GVHD Grau I (4,5%), 19 pacientes tiveram Grau II (21,3%), 14 tiveram Grau III (15,7%) e 03 tiveram Grau IV (3,3%).

Destes 40 pacientes, 20 (50%) pacientes foram submetidos a biópsia de tecido para confirmação diagnóstica de aGVHD, sendo 16 biópsias de pele, 03 intestinais e 01 hepática. As características clínicas dos pacientes que tiveram e não tiveram aGVHD são mostradas na Tabela 1 e os dados brutos encontram-se no anexo 2.

A probabilidade de sobrevida geral aos 12 meses foi 75% para o grupo de pacientes com ausência de aGVHD ou aGVHD Grau I contra 44% dos pacientes que tiveram aGVHD Grau II-IV e foi estatisticamente significativa ($p < 0,0001$).

Para testar o impacto de variáveis na sobrevida geral ao final de 12 meses, nós realizamos análise univariada e multivariada com as seguintes variáveis: GVHD agudo, GVHD crônico, risco da doença de base e regime de condicionamento, incompatibilidade ABO), cujos resultados são apresentados na tabela 2.

Na análise univariada, aGVHD ($p < 0,001$) e doença de alto risco ($p = 0,010$) foram fator de risco para menor sobrevida, enquanto GVHD crônico foi fator de risco para maior sobrevida ($p < 0,0001$) ao final de 12 meses de seguimento. A análise multivariada demonstrou que pacientes com diagnóstico de aGVHD tem maior risco de morrer ao final do primeiro ano ($p = 0,002$ Risco Relativo=3,2), enquanto que pacientes com diagnóstico de GVHD crônico tinham menor chance de morrer ao final do primeiro ano ($p < 0,0001$ Risco Relativo=0,17), demonstrando que aGVHD e crônico são fatores

independentes para menor e maior sobrevida ao final do primeiro ano, respectivamente.

Tabela 1 : Características Clínicas dos pacientes.

	aGVHD (N: 40)	Sem aGVHD (N: 49)	p
Mediana de idade	43(17-69)	37(16-63)	0,52
Incompatibilidade de sexo			
(F→M)	9	9	0,18
Risco da Doença			
Não Maligna	1	5	0,23
Baixo Risco	13	22	0,54
Intermediário	4	5	0,99
Alto Risco	22	17	0,25
Incompatibilidade ABO,			
sim	6	14	0,31
Profilaxia de GVHD			
CSA+ MTX	36	43	0,98
CSA+MMF	4	6	0,99
Tipo de transplante			
Mieloablativo	33	37	0,87
Toxicidade reduzida	7	12	0,61
Regime de			
Condicionamento			
Baseado em TBI	5	4	0,72
Baseado em Bussulfan	24	32	0,86
Outros	11	13	0,99
Fonte de Células			
Periféricas	32	34	0,74
Medula Óssea	8	15	0,48
Infecção Por CMV	16	13	0,39
Mediana de Enxertia			
Medular (dias)	14(10-20)	14(10-29)	0,30

Figura 8: Gráfico de sobrevida geral dos pacientes estratificados segundo a graduação de aGVHD.

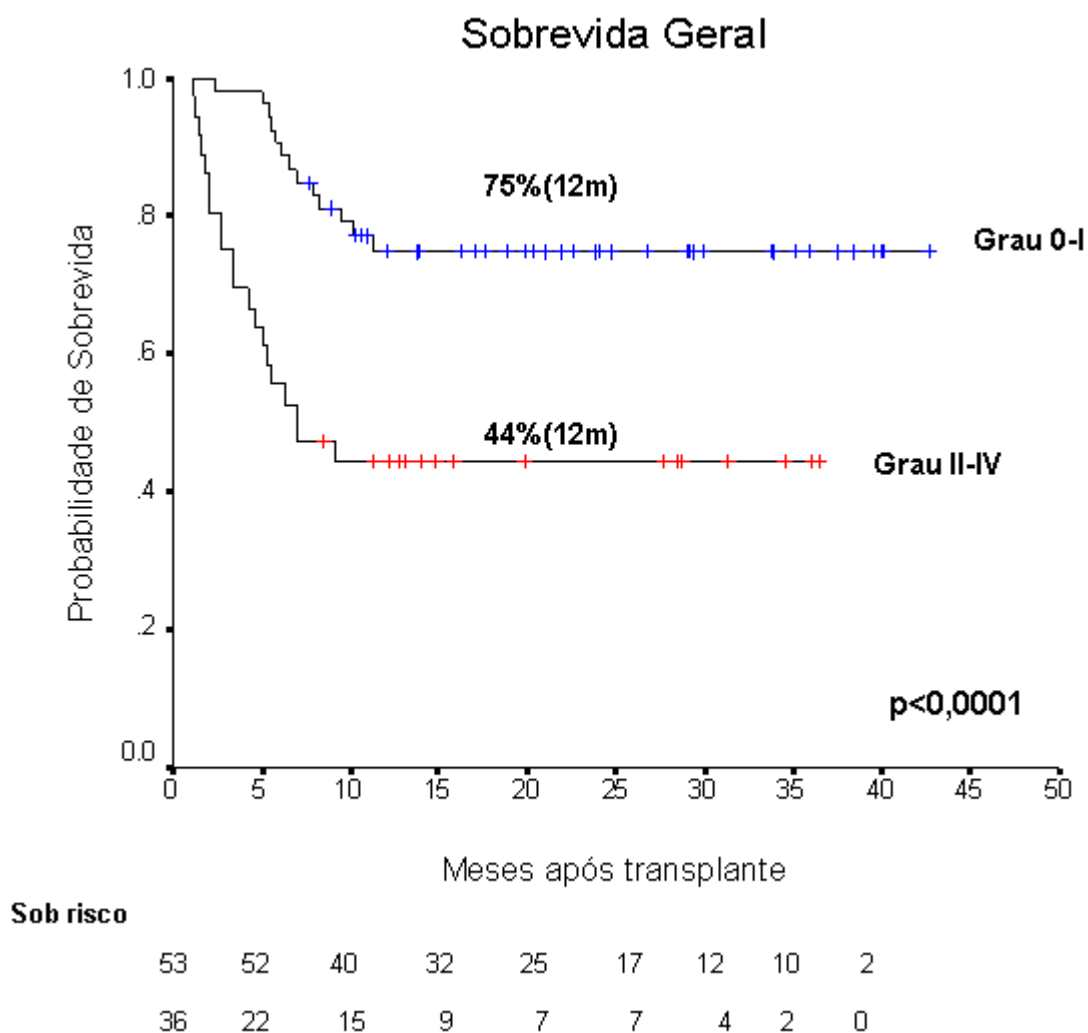


Tabela 2. Análise Univariada e Multivariada de Cox para a sobrevida geral ao final do primeiro ano de seguimento.

Fator de Risco	Análise univariada		Análise Multivariada	
	p	p	RR	IC (95%)
GVHD agudo II-IV vs Grau 0-I	<0,0001	0,002	3,2	1,5-6,7
GVHD crônico vs Não	<0,0001	<0,0001	0,17	0,08-0,36
Doença alto risco vs outros	0,010	0,56	----	----
Mieloablativo vs Toxicidade reduzida	0,55	----	----	----
Incompatibilidade ABO vs não	0,08	0,19	----	----

RR: risco relativo IC: intervalo de confiança

4.3 Identificação dos genes que compõem o preditor de aGVHD e teste do preditor

4.3.1. Identificação dos genes do preditor

Sabendo do impacto do diagnóstico de aGVHD na sobrevida geral, principalmente quando estão presentes estádios clínicos de moderado a grave (grau II-IV)⁽⁶⁰⁾, surgimento de aGVHD hiperagudo⁽⁶¹⁾, ausência de resposta à terapia imunossupressora baseada em corticóides⁽⁶²⁾, nós buscamos um perfil de expressão gênica que pudesse prever, no momento da enxertia medular, quais os pacientes que evoluiriam para aGVHD ou não. Para responder futuramente outras perguntas de grande interesse clínico, como a capacidade de prever graduação ou refratariedade, nós buscamos primeiramente saber, neste estudo, a capacidade deste modelo em prever a ocorrência ou não da reação aloimune.

Para a análise, dividimos nosso grupo em treino (N:67 amostras) e teste (n:22 amostras). Utilizamos as amostras treino para construir o preditor e o aplicamos com validação cruzada em ambos os grupos.

Para a construção do preditor, dividimos conforme ilustração da figura 6, o grupo treino composto de 67 amostras em dois subgrupos dez vezes. Para cada divisão aleatória foi criado um subgrupo de 33 amostras (18 sem aGVHD e 15 com aGVHD) e outro com 34 amostras (19 sem aGVHD e 15 com aGVHD), mantendo sempre a proporção de pacientes com e sem aGVHD. Uma vez criados estes subgrupos, e sabendo que passaram pelos critérios de filtragem cerca de 3863 genes, utilizamos *recursive features elimination* com 1000 genes. Para cada subgrupo, geramos uma lista de genes com N genes e fizemos a intersecção desta lista entre cada dois subgrupos,

criando então, 10 listas de genes. Para a construção final do preditor, escolhemos genes que apareciam em pelo menos 50% das listas de genes.

Desta forma, o nosso preditor foi construído com 233 genes (anexo 3).

4.3.2. Teste do preditor

Como não há consenso sobre o melhor algoritmo matemático para predição de classes e o uso de múltiplos algoritmos aumenta a confiança e a validade dos resultados⁽⁵⁵⁾, usamos 7 diferentes métodos, assim como a validação cruzada do tipo “*leave-one-out*”.

Para verificar a capacidade de nosso classificador em prever aGVHD, nós o testamos no grupo treino com validação cruzada do tipo *leave-one-out*. Utilizando as 67 amostras, o preditor foi capaz de classificar 54 amostras (80,5%), sendo 24 amostras de aGVHD e 30 amostras de não-aGVHD. Treze amostras não puderam ser classificadas pelo nosso algoritmo por não estarem em concordância em pelo menos seis métodos estatísticos. Assim, nas 54 amostras, o preditor molecular teve sensibilidade de 87,5%, especificidade 93,3%, acurácia 90,7%, como demonstra a tabela 3.

Para que um teste diagnóstico tenha validade externa, é necessário que o mesmo forneça resultados comparáveis, quando aplicado por equipes diferentes, em ambientes diferentes, ou em pacientes diferentes, ou se os testes forem repetidos. Para testar a robustez e a aplicação geral de nosso preditor molecular composto por 233 genes, nós testamos em um grupo independente de 22 amostras, utilizando validação cruzada do tipo *leave-one-out*.

Nesta amostra independente, o preditor foi capaz de classificar 07 amostras de aGVHD e 09 amostras de não-aGVHD, com sensibilidade de 71%, especificidade de 78% com acurácia de 75%, como demonstra a tabela 3. Seis amostras não puderam ser classificadas por nosso algoritmo.

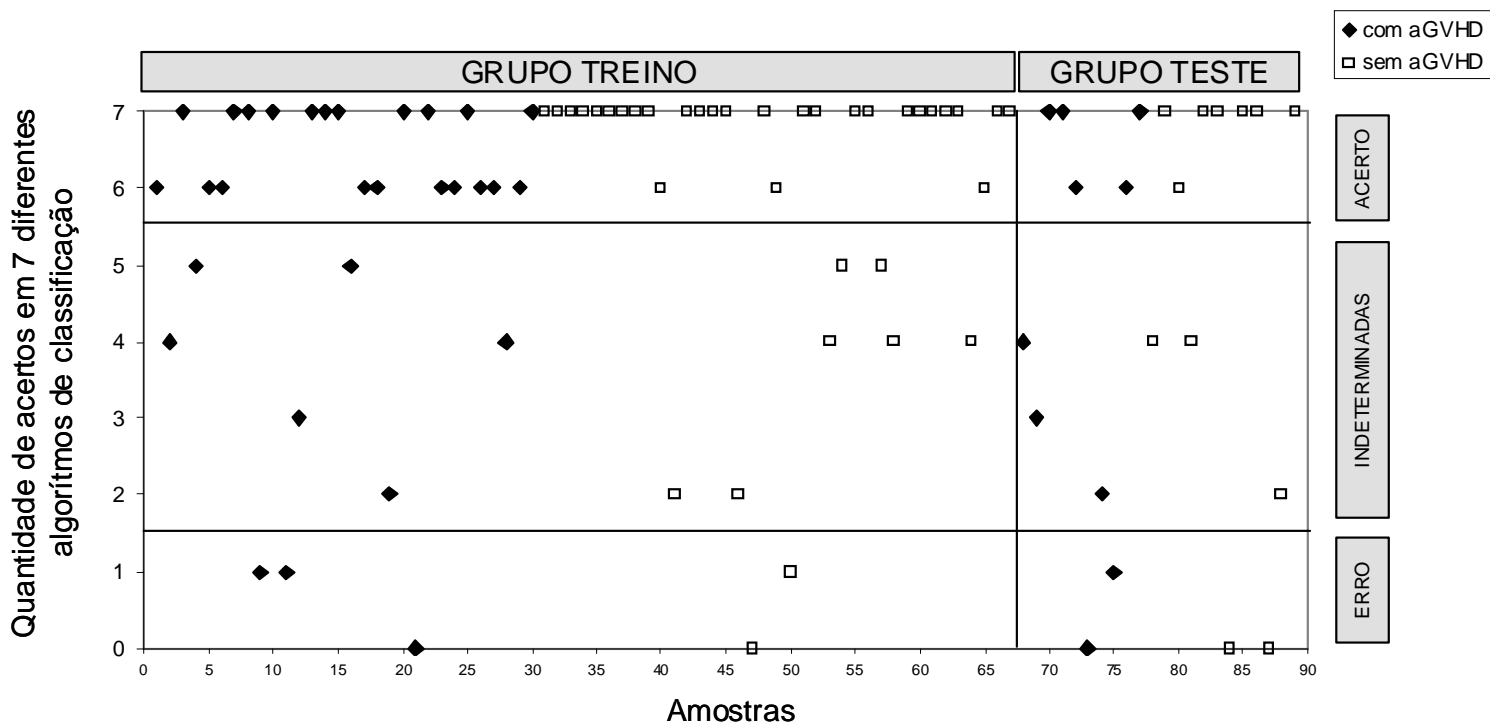
Com esta estratégia, 78,6% (N=70) das amostras puderam ser classificadas pelo nosso algoritmo e 19 amostras (13 amostras do grupo treino e 6 amostras do grupo teste) ficaram sem classificação.

Na figura 9 e nos anexos 4 e 5, podemos ver todas as amostras utilizadas neste estudo e a classificação das amostras de aGVHD e sem aGVHD no grupo treino e teste segundo acertos em cada um dos 07 algoritmos de predição.

Tabela 3: Desempenho do preditor no grupo treino e teste após exclusão de amostras inclassificáveis.

	GRUPO TREINO		GRUPO TESTE		
	aGVHD +	aGVHD -	aGVHD +	aGVHD -	
Acerto	21	28	5	7	
Erro	3	2	2	2	
Total	24	30	7	9	
Sensibilidade	Especificidade	Acurácia	Sensibilidade	Especificidade	Acurácia
87,5%	93,3%	90,7%	71%	78%	75%

Figura 9: Classificação de amostras de aGVHD e sem aGVHD no grupo treino e teste segundo acertos em cada um dos 07 algoritmos.



4.4 Análise dos genes diferencialmente expressos

Para melhor entender o papel funcional dos genes e a biologia dos eventos precoces que acontecem antes da emergência clínica de aGVHD, buscamos os genes diferencialmente expressos entre os dois grupos de pacientes. Utilizamos para isto uma ferramenta chamada comparação de classes.

A partir dos 233 genes presentes no preditor (anexo 3), nós aplicamos a comparação de classes, utilizando um limite de 10% para a taxa de descobertas falsas (*FDR – false discovery ratio*). Deste total de genes, obtivemos 70 genes, sendo que 54 estavam super-expressos no grupo com aGVHD e 16 sub-expressos no mesmo grupo como pode ser visto no anexo 6.

No anexo 6 consta uma tabela com os 70 genes diferencialmente expressos. Nesta tabela, a coluna que demonstra a relação entre as médias geométricas das duas classes significa que um valor menor que 1 indica expressão maior no grupo de aGVHD, enquanto um valor maior que 1 indica expressão maior no grupo sem aGVHD. Deste ponto em diante, todas as vezes que citarmos genes super ou sub-expressos estaremos nos referindo ao grupo com aGVHD.

Na figura 10 podemos observar um *heatmap* dos genes super e sub-expressos.

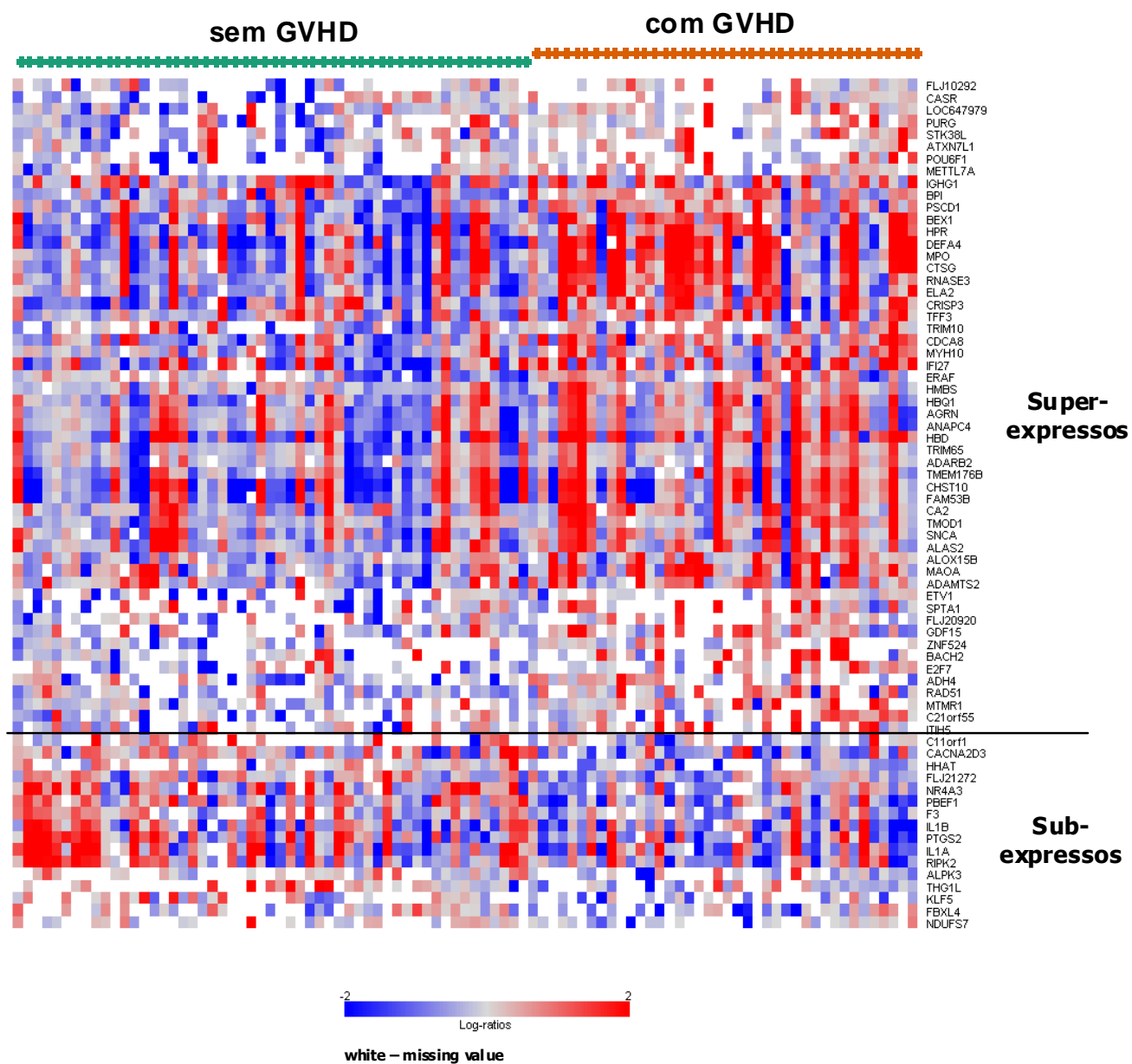


Figura 10: *Heatmap* dos genes diferencialmente expressos entre amostras dos pacientes que evoluíram ou não para aGVHD. Coluna representa amostras de pacientes e linha representa os genes. Os genes foram agrupados baseado no padrão de expressão de todas as amostras, utilizando *distância Euclidiana* *average linkage*. As amostras foram agrupadas de acordo com sua característica clínica e organizadas arbitrariamente dentro dos grupos.

4.5. Anotação funcional dos genes diferencialmente expressos

Classificação de genes baseado em termos do Gene Ontology (GO) é uma ferramenta atualmente usada para análise de listas de genes provenientes de experimentos de *microarray*. A análise automática de anotação funcional dos termos GO permite a identificação de famílias funcionais de genes que possam desempenhar papéis importantes relacionados a processos biológicos ou moleculares em determinados perfis de expressão⁽⁵⁹⁾.

O conjunto de genes diferencialmente expressos (70 genes do total de 233 genes) foi submetido a análise através do programa ExPlain *BIOBASE*, que realiza buscas em vários bancos de dados e fornece anotação funcional para cada grupo de genes.

Desta forma, fizemos uma busca utilizando o banco de dados do *Gene Ontology* (GO), o qual disponibiliza um vocabulário controlado para descrever cada gene e o produto dos genes para qualquer organismo. Os três parâmetros para a organização do banco GO são função molecular, processos biológicos e componente celular. Agrupando os genes baseados na similaridade de anotação funcional pode-se melhorar a interpretação biológica de listas de genes geradas pela técnica de *microarray*. O programa verifica quais os termos GO (chamados também de categorias) são estatisticamente enriquecidos na lista de interesse em comparação com o genoma total.

Os 70 genes diferencialmente expressos em pacientes com aGVHD foram divididos em dois grupos para análise funcional: 54 genes super-expressos e 16 genes sub-expressos no período precoce de aGVHD.

Esta análise demonstrou que o conjunto de genes super-expressos tinha 29 categorias funcionais enriquecidas (19 de processos biológicos e 10 de função molecular), sendo que alguns genes se encaixavam em mais de uma categoria, como pode ser visto nas tabelas 4 e 5.

Quanto à termos de processos biológicos, as principais categorias nos 54 genes super-expressos foram: resposta de defesa a bactéria, transporte de gases, metabolismo

de hemoglobina, hematopoese, transdução de sinal e metabolismo de neurotransmissores (Tabela 4).

Com relação à função molecular do grupo de 54 genes, os resultados consistem em: interação com bactéria, atividade de transportador de oxigênio, ligador de heme, constituinte estrutural do citoesqueleto, atividade de transdutor de sinal, atividade de peroxidase, atividade de oxi-redutase (Tabela 5).

Aqui, notamos que várias moléculas (*DEF4*, *RNASE3*, *CTSG*, *BPI*) pertenciam a um grupo de moléculas chamadas alarminas, recentemente descritas por Oppenheim e cols⁽⁶³⁾. Estas moléculas, liberadas diante de sinais de perigo, como lesão tecidual ou presença de microorganismos, serão discutidas mais detalhadamente, pois são responsáveis pela interação entre mecanismo inatos e adaptativos da resposta imune.

A análise dos 16 genes sub-expressos demonstrou o enriquecimento de 26 categorias (23 de processos biológicos e 03 de função molecular), sendo que alguns genes também se encaixavam em mais de uma classe, como pode ser visto nas tabelas 6 e 7.

Quanto às classes de processos biológicos, os 16 genes sub-expressos se agruparam em: geração de calor, regulação da proliferação celular, resposta inflamatória, resposta de defesa, formação de estrutura anatômica, regulação da atividade de fator de transcrição, regulação de ligação do DNA, regulação da cascata de I-kappaB kinase/NF-kappaB (tabela 6).

Quanto à classe função molecular, os 16 genes se agruparam em ligador do receptor de IL-1, atividade de doador de elétron e atividade de citocina (tabela 7).

Nós também checamos se dentro dos grupos de genes super e sub-expressos formavam-se redes de comunicações bioquímicas conforme dados da literatura. Para esta análise, nós utilizamos o software chamado Ingenuity Pathway Analysis⁽⁶⁴⁾. Nós não encontramos nenhuma rede representativa no grupo de genes super-expressos, mas muitos dos genes sub-expressos mostram-se conectados entre si na via do *NF-κB/p38 MAPK* (figura 11). Nesta via, *IL-1* via *RIPK2*, ativa *NF-κB* que ativa alguns genes-alvo encontrados em nossa lista (*F3*, *PTGS2*). Outros genes são também ativados pela *IL-1*

como *NAMPT*, *NR4A3*, *KLF5*. Podemos observar também um mecanismo de feedback nesta rede entre *IL-1* e *NF- κ B* e entre *PTGS2* e *NF- κ B*.

Tabela 4. Análise funcional do grupo de genes super-expressos segundo envolvimento em processos biológicos baseada no Gene Ontology (GO).

Identificador no GO	Símbolo do Gene	Termo no GO	Tamanho do grupo	Número de Genes	valor p
GO:0015669	CA2, GLOB-T1, HBD	transporte de gás	19	3	3.54E-05
GO:0020027	ALAS2, ERAF	metabolismo de Hemoglobina	6	2	0.00017
GO:0007009	AGRIN, MYH10	biogênese e organização da membrana plasmática	13	2	0.00087
GO:0006783	ALAS2, HMBS	biossíntese de Heme	14	2	0.00102
GO:0015671	GLOB-T1, HBD	transporte de oxigênio	16	2	0.00134
GO:0042168	ALAS2, HMBS	Metabolismo de Heme	18	2	0.00170
GO:0042417	MAOA, SNCA	Metabolismo de Dopamina	18	2	0.00170
GO:0006779	ALAS2, HMBS	Biossíntese de Porfirina	21	2	0.00232
GO:0033014	ALAS2, HMBS	Biossíntese Tetrapirrólico	21	2	0.00232
GO:0006778	ALAS2, HMBS	Metabolismo de Porfirina	28	2	0.00411
GO:0006584	MAOA, SNCA	Metabolismo Catecolaminas	29	2	0.00441
GO:0042742	BPI, DEFA4, RNASE3	Resposta de defesa a bactérias	99	3	0.00472
GO:0007165	AGRIN, CASR, GDF15, MYH10, PSCD1, STK38L	Transdução de Sinal	3996	6	0.00726
GO:0030097	ALAS2, ERAF, SPTA1, TRIM10	Hematopoese	225	4	0.00730
GO:0042133	MAOA, SNCA	Metabolismo de Neurotransmissor	38	2	0.00749
GO:0009617	BPI, DEFA4, RNASE3	Resposta a bactérias	118	3	0.00768
GO:0046148	ALAS2, HMBS	Biossíntese de pigmentos	40	2	0.00828
GO:0006952	BPI, CRISP3, DEFA4, ELA2, MPO, RNASE3, TFF3	Resposta de defesa	702	7	0.00956
GO:0048513	ADAMTS2, AGRIN, ALAS2, ALOX15B, CASR, ERAF, MYH10, POU6F1, SPTA1, TMEM176B, TRIM10	desenvolvimento de órgãos	1472	11	0.00985

Tabela 5. Análise funcional do grupo de genes super-expressos segundo sua função molecular baseada no Gene Ontology (GO).

Identificador no GO	Símbolo do Gene	Termo no GO	Tamanho do grupo	Número de Genes	valor p
GO:0008367	BPI, ELA2	Interação com bactérias	11	2	0.0005
GO:0005344	GLOB-T1, HBD	atividade de transportador de oxigênio	16	2	0.0011
GO:0020037	GLOB-T1, HBD, MPO, TMOD1	ligador de Heme	160	4	0.0015
GO:0004522	CASR, RNASE3	atividade de ribonuclease	21	2	0.0019
GO:0005200	AGRIN, MYH10, SPTA1	constituente estrutural do citoesqueleto	92	3	0.0029
GO:0004871	AGRIN, CASR	atividade de transdutor de sinal	2685	2	0.0065
GO:0051015	MYH10, SPTA1	ligador de filamento de actina	41	2	0.0071
GO:0004601	MPO, TMOD1	atividade de peroxidase	42	2	0.0074
GO:0016684	MPO, TMOD1	atividade de oxiredutase	42	2	0.0074
GO:0019825	GLOB-T1, HBD	ligador de oxigênio	42	2	0.0074

Tabela 6. Análise funcional do grupo de genes sub-expressos segundo envolvimento em processos biológicos baseada no Gene Ontology (GO).

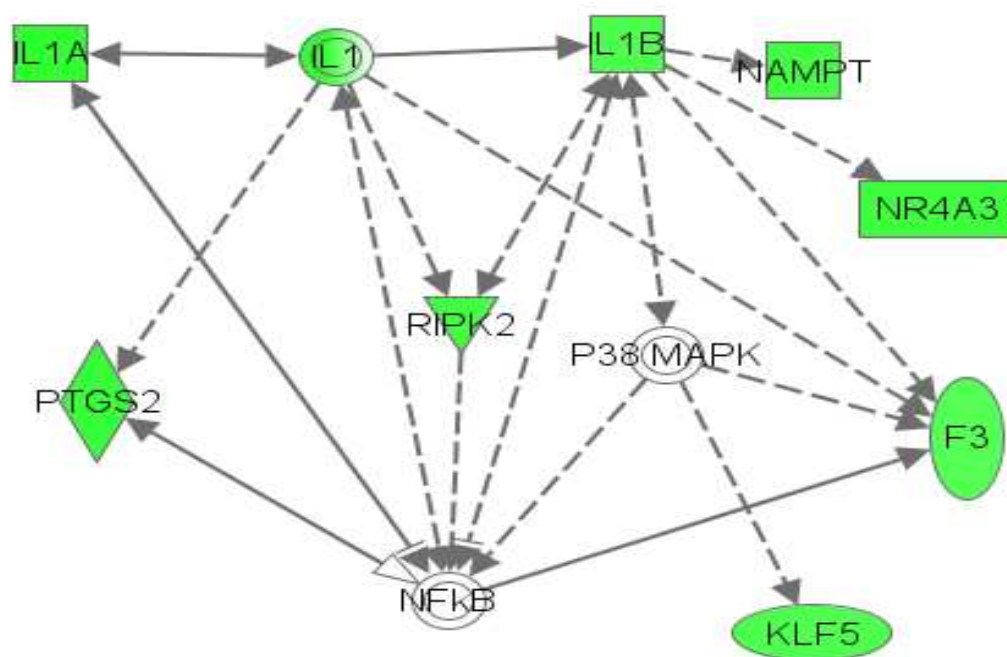
Identificador no GO	Símbolo do Gene	Termo no GO	Tamanho do grupo	Número de Genes	valor p
GO:0001660	IL-1beta, IL1A	Febre	2	3	3.09E-06
GO:0031649	IL-1beta, IL1A	Geração de calor	2	6	1.54E-05
GO:0008283	IL-1beta, IL1A, KLF5, PBEF1, PTGS2, RIPK2	Proliferação celular	6	869	0.00015
GO:0042127	IL-1beta, IL1A, KLF5, PBEF1, PTGS2	Regulação da proliferação celular	5	557	0.00019
GO:0009611	F3, IL-1beta, IL1A, PTGS2, RIPK2	Resposta a agressão	5	562	0.00020
GO:0006953	IL-1beta, IL1A	Resposta de fase aguda	2	32	0.00050
GO:0006954	IL-1beta, IL1A, PTGS2, RIPK2	Resposta inflamatória	4	375	0.00050
GO:0006952	FBXL4, IL-1beta, IL1A, PTGS2, RIPK2	Resposta de defesa	5	702	0.00056
GO:0051091	IL-1beta, RIPK2	Regulação positiva de atividade de fator de transcrição	2	36	0.00063
GO:0043388	IL-1beta, RIPK2	Regulação positiva de ligação ao DNA	2	38	0.00070
GO:0051099	IL-1beta, RIPK2	Regulação positiva de ligação	2	41	0.00082
GO:0048646	IL-1beta, KLF5, NR4A3	Formação de estrutura anatômica	3	197	0.00104
GO:0009605	F3, IL-1beta, IL1A, PTGS2, RIPK2	Resposta a estímulo externo	5	853	0.00136
GO:0051090	IL-1beta, RIPK2	Regulação da atividade de fator de transcrição	2	61	0.00181
GO:0000074	IL-1beta, IL1A, PTGS2	Regulação de progressão do ciclo celular	3	246	0.00197
GO:0051101	IL-1beta, RIPK2	regulação de ligação do DNA	2	70	0.00238
GO:0008285	IL-1beta, IL1A, PTGS2	Regulação negativa de proliferação celular	3	268	0.00251
GO:0051098	IL-1beta, RIPK2	Regulação de ligação	2	90	0.00391
GO:0043123	IL-1beta, RIPK2	Regulação positiva da cascata de I-kappaB kinase/NF-kappaB	2	103	0.00508
GO:0006950	F3, FBXL4, IL-1beta, IL1A, PTGS2, RIPK2	Resposta ao stress	6	1712	0.00563
GO:0043122	IL-1beta, RIPK2	Regulação da cascata de I-kappaB kinase/NF-kappaB	2	110	0.00578

GO:0002526	IL-1beta, IL1A	Resposta inflamatória aguda	2	123	0.00718
GO:0007249	IL-1beta, RIPK2	Cascata de I-kappaB kinase/NF-kappaB	2	145	0.00986

Tabela 7. Análise funcional do grupo de genes sub-expressos segundo sua função molecular baseada no Gene Ontology (GO).

Identificador no GO	Símbolo do Gene	Termo no GO	Tamanho do grupo	Número de Genes	valor p
GO:0005149	IL-1beta, IL1A	ligador do receptor de IL-1	2	15	9.47E-05
GO:0009053	NDUFS7, PTGS2	atividade de doador de elétron	2	18	0.00013
GO:0005125	IL-1beta, IL1A, PBEF1	atividade de citocina	3	257	0.00186

Figura 11. Rede de comunicação bioquímica da via NF-kB/p38 MAPK entre genes sub- expressos.



Genes coloridos são aqueles encontrados na nossa lista de genes sub-expressos;
linha sólida- interação/ligação direta; linha traçada – regulação de expressão

4.6. Agrupamento dos genes diferencialmente expressos segundo expressão em diferentes tipos celulares e tecidos humanos.

O objetivo desta análise foi tentar entender a origem celular dos genes encontrados diferencialmente expressos em nossos pacientes, uma vez que não usamos só um tipo determinado de células e vários tipos celulares poderiam contribuir para a expressão gênica. Para tal, nós utilizamos informações de expressão gênica obtidas de vários tecidos e tipos celulares de tecidos humanos publicados recentemente⁽⁵⁷⁾. Tecidos e linhagens celulares utilizadas nesta análise estão no site: http://wombat.gnf.org/samples/GeneAtlasv2_sample_info.html#human. Este trabalho realizado por Su e colaboradores explorou a expressão gênica de 49.859 transcritos em tecidos e linhagens celulares humanas e foi disponibilizado para uso da comunidade científica. Para esta análise, nós utilizamos somente tipos celulares e tecidos relacionados ao sangue e células do sistema imune.

Utilizando os símbolos gênicos, realizamos o mapeamento dos nossos 70 genes diferencialmente expressos em relação às informações de expressão obtidas pelo GNF SymAtlas.

Nesta análise, observamos 04 agrupamentos ou *clusters* de genes, sendo 03 grupos (grupo A,B, C) nos genes super-expressos e 01 *cluster* (grupo D) nos genes sub-expressos (Figuras 12 e 13).

Nós observamos que uma grande porção de genes super-expressos se agruparam predominantemente com células provenientes de medula óssea, com células de fígado fetal de 15-24 semanas e com células precursoras endoteliais e eritróide CD71, ambos provenientes de medula óssea (Cluster A). Este *cluster* foi composto pelo seguintes genes: *MYH10, FLJ10252, RAD51, CDCA8, TMOD1, SPTA1, HMDS, HBQ1, ERAF, HBD, ALAS2, TRIM10, BPI, MPO, ELA2, METTL7A, SNCA, CA2 E C21orf55*.

Também no grupo de genes super-expressos observamos dois clusters menores (B e C) associados a órgãos imunes periféricos e linfócitos T. O *cluster* B foi composto pelos seguintes genes: *CRISP3, IGHG1, IFI27, MAO, TMEM176B e HPR*. Os tipos de

tecido agrupados foram: tonsilas, timo e linfonodos. O *cluster* C foi composto dos seguintes genes: *MTMR1*, *POU6F1*, *ETV1*, *PSCD1*, *FAM53B* e *BACH2*. As linhagens celulares com os quais os genes se agruparam foram: células T CD4+ e CD8+ provenientes de sangue periférico.

No grupo de genes sub-expressos, predominou a assinatura molecular de células de sangue periférico (Cluster D). Este *cluster* foi composto pelo seguintes genes: *PTSG2*, *RIPK2*, *PBEF1* (*NAMPT*), *CACNA2D3* e *IL1B*. Os tecidos e linhagens celulares foram: sangue total, células dendríticas BDCA4+ de sangue periférico, células CD33+ (mielóide) de medula óssea, células CD14+ (monócitos) de sangue periférico e células CD34+ de medula óssea.

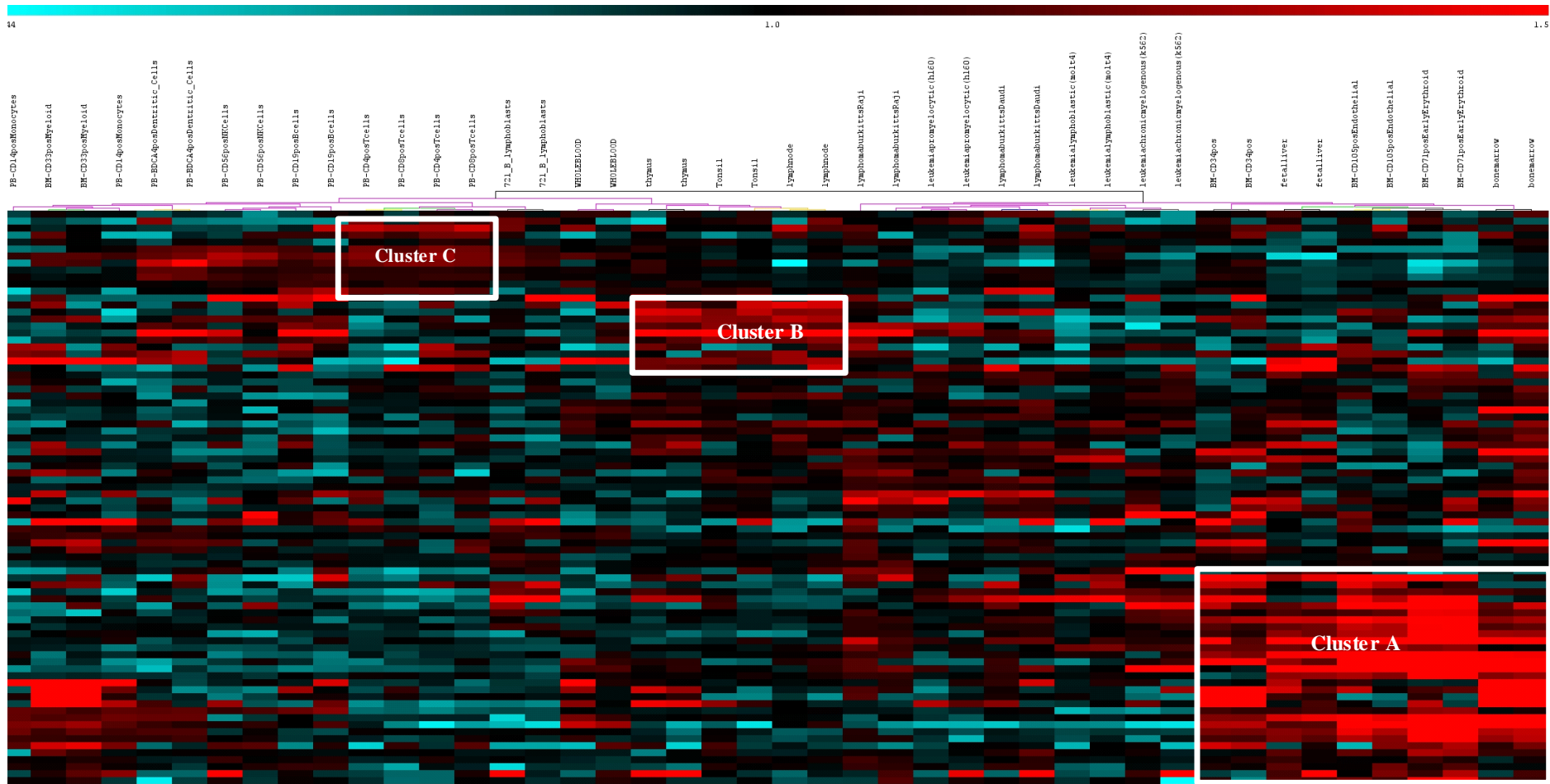


Figura 12. Heat Map dos 54 genes super-expressos e o agrupamento dos mesmos em relação a expressão deste genes em diversos tecidos e linhagens celulares humanas obtidas do SymAtlas. Os agrupamentos A,B e C demonstram um conjunto de genes super-expressos. Colunas representam tecidos e linhas representam os genes.

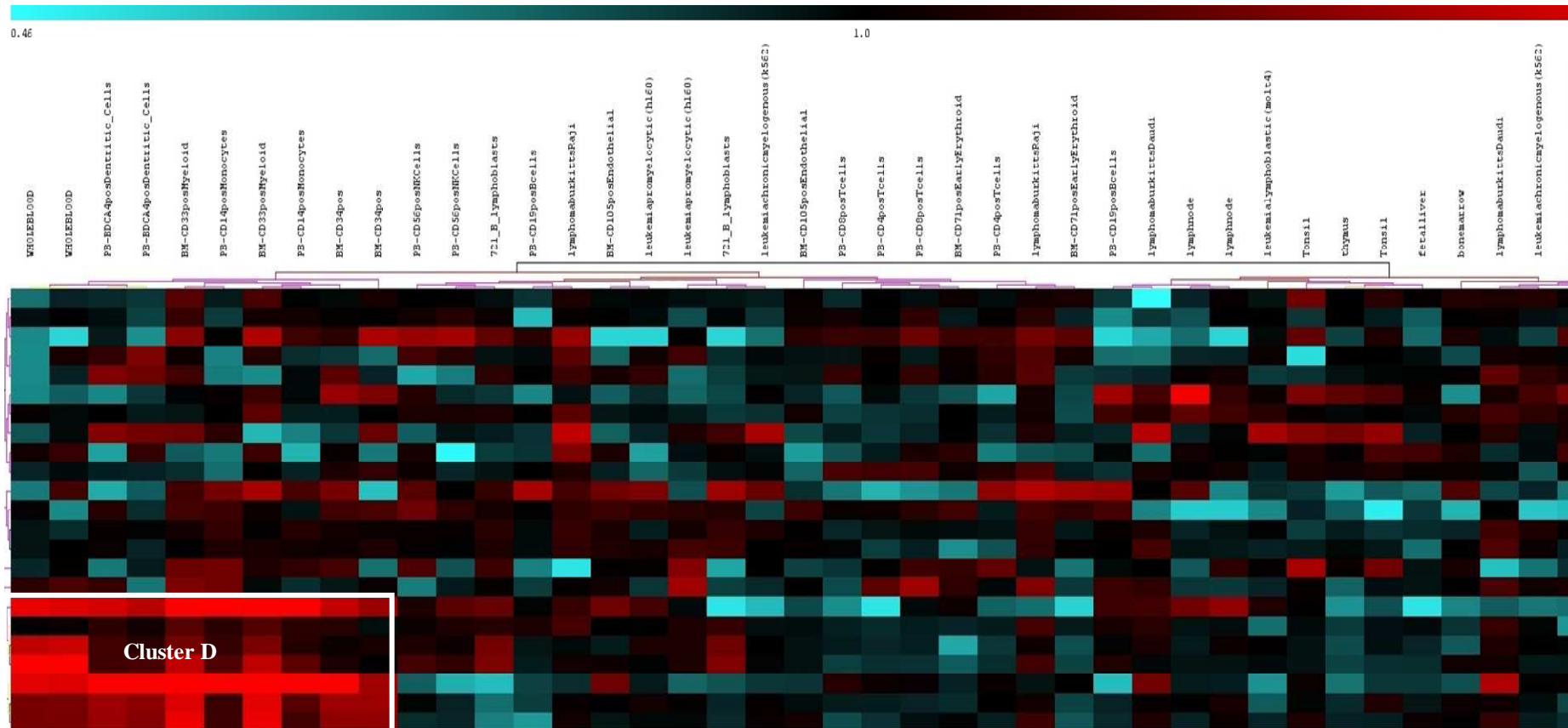


Figura 13. Heat Map dos 16 genes sub-expressos e o agrupamento dos mesmos em relação a expressão deste genes em diversos tecidos e linhagens celulares humanas obtidas do SymAtlas. O agrupamento D demonstra um conjunto de genes sub-expressos. Colunas representam tecidos e linhas representam os genes.

5. Discussão

5.1. Predição de aGVHD como ferramenta diagnóstica

Quimioterapia associada ou não a radioterapia, seguida de infusão de células tronco hematopoéticas é terapia curativa em várias doenças hematológicas. Células do doador constituem uma fonte rica em células progenitoras hematopoéticas capazes de reconstituir a hematopoese normal, embora sejam também ricas em células T CD4+ e CD8+. Além de facilitarem a enxertia medular e desempenharem potencial efeito anti-tumor, estas células desencadeiam a doença do enxerto contra o hospedeiro aguda^(65,66).

Doença do enxerto versus hospedeiro é a complicação mais frequente após transplante alogênico de células tronco hematopoéticas. A emergência de aGVHD é associada ao aumento da morbidade e mortalidade, particularmente em pacientes que não respondem à terapia inicial com corticóides. A intensidade da aGVHD é clinicamente graduada de I-IV (Gluksberg), avaliando os três principais órgãos-alvo e a gravidade da doença é preditiva de maior morbi-mortalidade^(4,67). aGVHD pode ser a causa direta da morte através de lesão de órgão-alvo ou predispondo o paciente à imunossupressão intensa e ao desenvolvimento de infecções graves. Embora aGVHD tenha um importante papel em erradicar células malignas, GVHD permanece como maior determinante de sobrevida, morbidade pós transplante e de qualidade de vida dos pacientes, especialmente quando longa imunossupressão é requerida para controle desta complicação^(8,60,68). O impacto destes efeitos na sobrevida depende da gravidade da aGVHD, sensibilidade ao tratamento imunossupressor e o sucesso do efeito enxerto versus malignidade⁽⁶⁹⁾.

A presença de aGVHD de moderado a grave *per si*, confere prognóstico adverso aos pacientes, pois metade deles não responderão à terapia tradicional baseada em corticóides, e é por este motivo que se torna essencial o desenvolvimento de ferramentas capazes de prever e entender melhor os mecanismos biológicos que culminam com este tipo de resposta imune. Diagnóstico precoce de aGVHD pode

aumentar a segurança dos TCTH alogênicos, permitindo intervenção necessária para alguns, e evitando imunossupressão exagerada para outros.

Neste estudo com 89 pacientes, 40(45%) desenvolveram GVHD agudo grau I-IV. Estes dados não são diferentes de dados previamente publicados quanto à incidência de GVHD agudo⁽⁷⁰⁾, assim como o impacto de GVHD crônico na sobrevida geral⁽⁷¹⁾. Nós observamos, ao final do primeiro ano, uma menor sobrevida geral em pacientes que desenvolveram aGVHD de moderado a grave (grau II-IV) quando comparado com pacientes que tiveram aGVHD grau 0-I. Utilizando o critério de Glucksberg modificado, Socié e cols demonstraram que a sobrevida geral, ao final do primeiro ano, decresce de acordo com a intensidade de aGVHD, sendo >90% para grau I, 80% para grau II, 30% para grau III e <10% para grau IV⁽⁷²⁾. Embora sejam resultados concordantes, o impacto da ocorrência de aGVHD na sobrevida geral, principalmente quando presente aGVHD grau I-II, permanece controverso^(73,74). Isso ocorre devido a uma combinação complexa do potencial efeito positivo do enxerto versus tumor exercido pela reação aloimune, com os efeitos deletérios secundários a lesões de órgãos acometidos por aGVHD, infecções e outras complicações decorrentes de aGVHD, causando efeitos opostos na sobrevida geral.

A técnica de microarray permitiu a descoberta de marcadores biológicos em diversas doenças. Enquanto o padrão de expressão gênica diferencial promove descobertas acerca da origem de doenças⁽⁷⁵⁾, promove também o desenvolvimento de preditores moleculares para diversas situações clínicas^(76,77). O desenvolvimento de biomarcadores úteis em diagnóstico ou prognóstico demonstra a possibilidade de uma medicina e tratamento personalizado.

Um ponto interessante e importante a ser abordado neste tipo de estratégia é que quando utilizamos a ferramenta de microarray para classificar eventos biológicos já instalados, os resultados, em geral, são melhores^(35,78) do que quando desejamos prever um evento com considerável antecedência^(79,80). Nesta última situação, devemos levar em consideração que o evento estudado possa ser uma medida dependente do tempo, e no momento em que o teste é aplicado, o processo biológico pode ainda não

ter se instalado. Deste modo, a predição de eventos futuros, em geral, cursa com valores de acurácia mais modestos.

Em TCTH, fatores de risco atualmente reconhecidos pela literatura não são capazes de fornecer uma estimativa precisa do risco individual de aGVHD e dessa forma não permitem a individualização da terapia. Até o presente momento, não existe nenhum teste sérico validado para diagnóstico de aGVHD, embora várias proteínas séricas tenham sido descritas como biomarcadores de aGVHD em alguns estudos⁽⁸¹⁻⁸⁴⁾. A habilidade de prever estes pacientes poderá ter grande aplicabilidade clínica na era das múltiplas opções terapêuticas.

A possibilidade de prever a ocorrência ou não de aGVHD utilizando uma amostra de sangue periférico colhida durante a enxertia medular (pelo menos uma semana de antecedência do evento) pode ser bastante útil do ponto de vista da intervenção terapêutica. Desta forma, testamos a capacidade de nosso preditor tanto no grupo treino quanto no grupo teste, utilizando a validação cruzada do tipo *leave-one – out*. Interessante perceber que testamos nosso preditor primeiro no grupo treino, pois seria importante primeiro verificar se existia uma boa acurácia neste grupo, para só depois analisar o grupo teste. Assim, dos 89 pacientes, foi possível fazer predição de presença ou ausência de aGVHD na maior parte de nossa população, 70 pacientes (78,6%). Em 19 (21,4%) pacientes, nosso preditor não conseguiu classificar a ocorrência ou não de evento e isso pode ter ocorrido devido a duas possíveis causas: primeiro, o fenômeno aloimune ainda não foi iniciado e portanto não pode ser detectado, segundo, algum processo biológico subclínico causando perturbação suficiente para não permitir a correta classificação.

Desta forma, nós utilizamos um paradigma novo para métodos diagnósticos que permite, além de discriminar entre pacientes com e sem doença, definir um grupo separado dos pacientes em que o classificador não se aplica (i.e. não é capaz de dar uma resposta suficientemente precisa). Esta solução foi recentemente proposta e utilizada com sucesso para a discriminação entre subtipos de linfoma de grandes células B⁽⁸⁵⁾.

Como criamos nosso preditor em um momento no qual nenhum sinal clínico de aGVHD estava presente, obtivemos uma predição relativamente modesta. Ainda assim, não conhecemos ferramenta disponível na literatura que permita, com uma acurácia de 75 %, predizer a ocorrência de aGVHD, na maior parte da população, com uma antecedência de 1 semana.

Recentemente, um trabalho publicado por Baron e cols⁽³⁹⁾, descreveu a predição de GVHD (agudo e crônico) com uma acurácia de 65% e de GVHD crônico em torno de 77%, baseado no perfil de expressão nas células do doador, obtida previamente à coleta de células tronco hematopoéticas. Neste estudo, ainda que o perfil de expressão das células T do doador possa refletir seu status de ativação imune, isto pode não estar relacionado à capacidade do doador em montar uma resposta do enxerto contra o hospedeiro após o contato com aloantígenos. Nós acreditamos que o momento da enxertia medular seja um bom momento para quantificar as mudanças de expressão gênica que possam predizer a ocorrência de aGVHD, pois invariavelmente, aGVHD inicia-se após este período e contagens de subpopulações de células, durante esta fase, tem sido relacionadas a aGVHD⁽⁸⁶⁾. Além disso, acreditamos que a exposição de células do sistema imune a agentes imunossupressores, habitualmente utilizadas após transplante, possa conferir mudanças na expressão gênica que não seriam observadas nos doadores.

A semelhança dos dados demográficos e o impacto destas variáveis na sobrevivência pós transplante nos sugerem uma possível validade externa de nosso preditor para aGVHD, realidade essa a ser confirmada em um grupo independente de pacientes de outros centros.

Desta forma, neste estudo, nós criamos um preditor molecular para aGVHD, contendo 233 genes, capaz de predizer, durante a recuperação medular, a ocorrência de doença do enxerto aguda em um grupo independente de 22 pacientes com uma acurácia de 75%, sensibilidade de 71% e especificidade de 78%.

5.2 . Microarrays e a descoberta de novos eventos biológicos envolvendo aGVHD

A técnica de microarray representa uma importante ferramenta para análise de expressão gênica global. Com ela, a ciência é capaz de determinar a interação entre milhares de genes simultaneamente, em estados de saúde ou patológicos. Ela permite, além de predição de eventos, comparar expressão de tecidos normais e doentes, encontrar novos biomarcadores para terapia alvo, determinar resistência à terapia e separar pacientes de pior evolução clínica em determinados cânceres. A análise simultânea de milhares de genes permite observar a interação de genes já descritos na literatura, mas também possibilita a descoberta de novos genes ou vias em atividade que não eram descritas. Essa perspectiva de descobrir novos fenômenos biológicos gera uma infinidade de conhecimento, que abrem portas para novas pesquisas e novas descobertas.

A partir da análise dos genes diferencialmente expressos em pacientes que evoluíram com aGVHD e suas categorias funcionais, observamos o enriquecimento de várias categorias gênicas funcionais, incluindo genes relacionados à resposta à presença de bactérias (tabela 4). Chamou-nos a atenção dentre estes genes, aqueles recentemente denominados na literatura como alarminas⁽⁶³⁾.

Nossa lista de genes super-expressos possui 4 representantes das alarminas, sendo eles: Alfa -Defensina 4 (DEFA4), RNase 3 (proteína catiônica eosinofílica), Catepsina G (CTSG) e BPI (Bactericidal/Permeability-Increasing protein).

Diante de “sinais de perigo”, definidos como a presença de lesão tissular ou mediadores inflamatórios intercelulares, o organismo produz e libera mediadores que alertam o sistema imune inato e adaptativo para a ativação de mecanismos de defesa e reparo⁽⁸⁷⁾. “Sinais de perigo” expressos por tecidos lesados são críticos para a ativação de células dendríticas do hospedeiro e são necessários para a iniciação da resposta primária e secundária. Este conceito explica um grande número de observações clínicas em transplante como o aumento do risco de aGVHD em pacientes com estágios

avançados de leucemias agudas, certos regimes de condicionamento e história de infecções virais^(88,89).

Conforme proposto por Oppenheim e cols⁽⁶³⁾, alguns destes mediadores endógenos da resposta inata chamados alarminas têm como características serem liberados em resposta à lesão tecidual ou infecção bacteriana, possuírem atividade quimiotática e promoverem ativação de células apresentadoras de antígenos (APC).

As defensinas alfa consistem em uma família de pequenos peptídeos antimicrobianos, sendo também conhecidas como peptídeos humanos neutrofílicos. São constitutivas de neutrófilos, correspondendo a 70% do conteúdo de grânulos azurófilos⁽⁹⁰⁾. São descritas quatro alfa-defensinas (DEF-1-4) em humanos, produzidas por neutrófilos e monócitos^(90,91) e duas alfa-defensinas (DEF-5-6) produzidas pelas células de Paneth do intestino delgado, células epiteliais de tecido placentário, nasal e brônquico⁽⁹³⁾. Além da atividade antimicrobiana, tem sido atribuído às alfa defensinas a capacidade de exercerem efeitos quimiotáticos sobre células T naive CD4+CD45RA, CD8+ e células dendríticas imaturas^(94,95). A busca de novas atividades biológicas das defensinas levou à observação de atividade quimiotática seletiva para células dendríticas imaturas provenientes de células progenitoras hematopoéticas CD34 + ou de monócitos de sangue periférico^(96,97). Por suas características, presumivelmente, as defensinas devem estar presentes no sítio de inflamação, formando gradiente quimiotático para células dendríticas imaturas. E, sabendo que células dendríticas são cruciais para a indução de resposta adaptativa, é esperado que o grande recrutamento de células dendríticas para o sítio de deposição antigênica induza a uma grande resposta imune antígeno-específica.

Catepsina G é uma serino proteinase primariamente derivada de grânulos azurófilos de neutrófilos, mas também encontrada em monócitos e mastócitos. Tem atividade antimicrobiana, promove degradação de matriz extracelular, promove vasorregulação⁽⁹⁸⁾, ativação de elastase neutrofílica⁽⁹⁹⁾, processamento de IL-8⁽¹⁰⁰⁾ e efeito quimiotático sobre monócitos⁽¹⁰¹⁾. Adicionalmente, a Catepsina G foi descrita como agente mitogênico de células T em camundongos e com capacidade de se ligar e

promover ativação em linfócitos B, linfócitos CD4+, CD8+ e células NK de humanos^(102,103).

A cathepsina G é responsável pela transformação de IL-8 em sua forma mais ativa em tecidos acometidos por processo inflamatório⁽¹⁰⁰⁾. IL-8 é uma citocina pró-inflamatória liberada durante a fase aguda de infecção, com reconhecido poder quimiotático e de ativação para neutrófilos, exibindo também uma fraca ação em basófilos⁽¹⁰⁴⁾, e também descrita como um citocina de fase tardia em rejeição de órgãos sólidos, principalmente em rejeições irreversíveis⁽¹⁰⁵⁾. Níveis séricos elevados de IL-8 vem sendo descrito em pacientes com emergência de aGVHD^(78,106), e é possível que por sua potente ação quimiotática, possa auxiliar na migração de células efetoras para órgãos alvos de aGVHD⁽¹⁰⁷⁾. Desta forma, a cathepsina G pode servir como um sinal derivado de neutrófilos que induz subsequente resposta imune dependente de células T e monócitos.

As células da imunidade inata formam a primeira linha de defesa do organismo e participam de eventos que iniciam a resposta imune a um aloenxerto⁽¹⁰⁸⁾. O papel biológico dos eosinófilos na patogênese de aGVHD é menos conhecido do que seu papel na alergia e na asma. Evidências recentes da literatura têm estabelecido um novo paradigma para a função dos eosinófilos durante a inflamação, para além de seu papel como célula efetora com atividade citotóxica não-específica. Por exemplo, Lofti e cols demonstraram que os eosinófilos são capazes de induzir maturação de células dendríticas, e que isso poderia contribuir para a regulação da imunidade no cenário de lesão tecidual associada a inflamação crônica que ocorre na rejeição de órgãos sólidos⁽¹⁰⁹⁾. Outras evidências têm implicado os eosinófilos como parte dos processos de remodelamento e de reparo iniciados com a resposta inicial ao dano tissular e a necrose celular⁽¹¹⁰⁾. Recentemente, um trabalho correlacionou a gravidade de aGVHD intestinal com a presença de eosinófilos, ativação e aumento da densidade grânulos citoplasmáticos, sugerindo um recrutamento destas células para o tubo gastrointestinal⁽¹¹¹⁾.

Entre os genes super-expressos, podemos destacar a presença de RNase3 ou proteína catiônica eosinofílica. Esta proteína tem como funções principais, capacidade citotóxica, atividade imunomoduladora, regulação da atividade fibroblástica e indução de secreção em mucosa aérea⁽¹¹²⁾. No contexto de rejeição do enxerto de rins e fígado transplantados, eosinófilos tissulares mostram sinais de ativação com depósito de proteína catiônica⁽¹¹³⁾ e sua densidade tem valor preditivo para rejeição resistente a corticóide^(114,115). Este papel expandido dos eosinófilos além de estabelecer uma ligação mais forte entre as células da imunidade inata e a resposta imune adaptativa, ainda pode fornecer provável explicação para a presença desta células em tecidos acometidos por aGVHD.

Em resumo, o aumento da expressão de transcritos de alarminas durante a enxertia medular, provavelmente liberadas de tecidos do receptor, pode potencialmente modular a resposta imune de células T provenientes do doador, através do recrutamento destas células por indução de sinal quimiotático para os tecidos lesados do receptor, ou por sua ativação através de sinais co-estimulatórios, levando à ativação e proliferação das células T do doador. Esta proliferação expande a alorreatividade do pool de células T do doador, causando uma lesão imuno-mediada, levando a mais liberação de alarminas. Desta forma, as alarminas liberadas após lesão tissular devem atuar alterando e intensificando a resposta aloimune.

A integridade do tubo gastrointestinal tem um papel importante para o desencadeamento de aGVHD. Irradiação corporal total é particularmente importante porque induz apoptose de células epiteliais do tubo gastrointestinal⁽¹¹⁶⁾, permitindo que produtos microbianos imuno-estimulatórios como o lipo-polissacáride (LPS), entrem na circulação sistêmica, levando à amplificação de aGVHD⁽¹¹⁷⁾. A produção de alfa defensinas, BPI e catepsina pode ser fortemente induzida em resposta a estímulos pro-inflamatórios como LPS⁽¹¹⁸⁾.

O aumento da expressão de alarminas e de outras moléculas como MPO e Elastase no grupo que evoluiu com aGVHD, pode estar relacionado com a lesão tissular induzida pelo regime de condicionamento e liberação de LPS por bactérias gram

negativas intestinais promovendo recrutamento de neutrófilos. Esta é uma hipótese que pode ser verificada através da medida de LPS no soro destes pacientes.

A presença de MPO e ELA2 na lista de genes super-expressos, que tem níveis maiores em células precursoras mielóides do que em neutrófilos periféricos⁽¹¹⁹⁾, associados ao aumento de expressão de genes ditos alarminas, numa situação em que a separação de células linfomononucleares foi realizada, nos leva a crer que estes genes possam estar aumentados em suas atividades, já em precursores, determinando a resposta imune. O aumento da atividade antimicrobiana em resposta à presença de LPS, deve influenciar o aumento da expressão destas moléculas precocemente, iniciando a ativação do sistema imune inato e adaptativo.

Esta hipótese ganha força quando observamos o cluster A, composto pelos seguintes genes: *MYH10*, *FLJ10252*, *RAD51*, *CDCA8*, *TMOD1*, *SPTA1*, *HMDS*, *HBQ1*, *ERAF*, *HBD*, *ALAS2*, *TRIM10*, *BPI*, *MPO*, *ELA2*, *METTL7A*, *SNCA*, *CA2* e *C21orf55* e que se agruparam com os seguintes tipos celulares: células de medula óssea, células CD34+ de medula óssea, fígado fetal de 15-24 semanas, células endoteliais provenientes de medula óssea, precursor eritroide CD71 obtido também de medula óssea. Estes dados são muito interessantes por dois aspectos: primeiro, podemos observar que a expressão deste grupo de genes estudados durante a enxertia medular de pacientes transplantados possui expressão semelhante à expressão de células bastante imaturas. Podemos observar o agrupamento de nossos genes com células CD34 de medula óssea e células progenitoras hematopoéticas provenientes de fígado fetal entre 15-24 semanas. Isto sugere circulação de células precursoras CD34 + durante a enxertia medular após transplante. É descrito que células CD34+ co-expressam em sua superfície celular moléculas de CD18 e CD86 constitutivamente, e podem apresentar aloantígenos eficientemente para células T. Por outro lado, células CD34+/CD18- e CD34+/CD86- possuem fraca atividade como células apresentadoras de antígeno^(120,121). Estes dados são consistentes com a existência de subtipos celulares precursores com atividades distintas de apresentação antigênica.

Notamos um enriquecimento de genes relacionados à hematopoese no grupo de pacientes que evoluíram com aGVHD, principalmente relacionados à produção de hemoglobina e transporte de gases. O aumento de expressão destes grupos de genes em pacientes com aGVHD não pode ser explicado pelo aumento da hematopoese neste período pós transplante, mas provavelmente, pela lesão endotelial que se segue ao TCTH. É conhecida a relação entre o aumento da incidência de toxicidade relacionada ao regime de condicionamento e a incidência de aGVHD⁽¹²²⁾. Evidências recentes demonstram que o endotélio pode ser um alvo para aGVHD⁽¹²³⁾. Durante a enxertia medular, o endotélio vascular do hospedeiro provavelmente deve ser o primeiro tipo celular com o qual os linfócitos T do doador têm contato após TCTH, potencialmente induzindo aGVHD. A diversidade de apresentação clínica de lesão endotelial após TCTH como microangiopatia trombótica⁽¹²⁴⁾, doença veno-oclusiva⁽¹²⁵⁾, síndrome de leak capilar⁽¹²⁶⁾, síndrome da enxertia⁽¹²⁷⁾ e hemorragia alveolar difusa⁽¹²⁸⁾ e a dificuldade de separá-los de aGVHD em muitas situações, leva a crer que aGVHD endotelial possa ser um elemento essencial nestas várias síndromes de lesão endotelial.

Recentemente foi demonstrado que a detecção de células endoteliais circulantes, provenientes de medula óssea, é sensível e específica como sinal de lesão endotelial, além do que um maior número de células endoteliais foi detectado após regimes de condicionamento mieloablativos quando comparadas com regimes de toxicidade reduzida⁽¹²⁹⁾. Interessante observar que a resposta a esse dano endotelial é a liberação de grande quantidade de eritropoetina, estimulando a medula óssea a aumentar o número de células endoteliais precursoras⁽¹³⁰⁾, na tentativa de reparar o dano. Esse aumento da produção de eritropoetina em resposta à lesão endotelial, possivelmente deve desviar e aumentar a atividade do maquinário celular de precursores eritróides para o processo de produção e metabolismo de hemoglobina.

Foi observado também uma expressão aumentada do gene *IFI27* (Interferon, alpha-inducible protein 27), também chamado de *P27* ou *ISG12*, em pacientes com aGVHD. Este gene faz parte de um conjunto de centenas de genes que são induzidos por *IFN*⁽¹³¹⁾. Interferon é uma família de citocinas com uma variedade de funções

biológicas, incluindo efeito anti-tumoral, atividade imunomodulatória e antiparasitária⁽¹³²⁾. Resposta biológica ao interferon é iniciada pela ligação do IFN a receptores específicos de superfície celular e subsequente ativação de vias de transdução de sinais e manifestada através da expressão de um grande número de genes induzidos por IFN, sendo estes os responsáveis pela ação biológica. Encontrar genes induzidos por IFN nesta fase do transplante está de acordo com dados recentemente publicados⁽¹³³⁾. Em um modelo de transplante murino⁽¹³³⁾, verificou-se precoce aumento da expressão de genes induzidos por *IFN- γ* , por volta do sétimo dia após o transplante, e antes da evidência histológica de aGVHD.

Em muitos modelos de aGVHD experimental, células T produzem grandes quantidades de *IFN- γ* ^(5,134, 135). Estes trabalhos sugerem que o *IFN- γ* esteja envolvido em muitos passos da fisiopatologia de aGVHD, atuando no aumento da expressão de moléculas de adesão, quimiocinas, HLA e moléculas envolvidas na apresentação antigênica. Desta forma, o interferon facilitaria a apresentação antigênica e promoveria recrutamento de efetores para órgãos-alvo de GVHD⁽¹³⁶⁾. Além disso, *IFN- γ* aumenta a expressão de Fas em tecidos-alvo, sugerindo que citotoxicidade de células T mediada por Fas seja essencial para o desenvolvimento de GVHD^(137,138).

Um estudo recente comparou a expressão diferencial de genes durante aGVHD e após o controle do fenômeno imunológico em oito pacientes submetidos a transplante de cordão umbilical. Neste estudo, foi observado o aumento da expressão de *IFI27* em monócitos, células T CD4+ e CD8+ durante aGVHD, sugerindo que esta molécula possa desempenhar importante papel em diferentes subpopulações na fisiopatologia de aGVHD⁽¹³⁹⁾.

O *cluster B* foi composto por genes que se agruparam com os seguintes tipos de tecido: tonsilas, timo e linfonodos, enquanto o *cluster C* foi composto de genes que se agruparam com células T CD4+ e CD8+ provenientes de sangue periférico. Embora os genes encontrados não pareçam ter um papel conhecido em aGVHD, e portanto, possam ser alvos de próximos estudos, esta observação de clusterização com linfócitos periféricos, órgãos linfóides primários e secundários é esperada e deve estar

relacionada ao tráfego de linfócitos que ocorre precocemente após o transplante. Linfócitos do doador, após poucas horas da infusão das células alogênicas, se acumulam em órgãos linfóides periféricos como linfonodos, placa de Peyer e baço. Nestes locais, as células T alorreativas sofrem expansão, diferenciando-se em efetores e produzindo *IFN- γ* , *TNF- α* e outras citocinas. Mediadores inflamatórios e produtos bacterianos (LPS) em sinergia com *IFN- γ* , induzem à produção de quimiocinas induzidas pelo *IFN- γ* por células apresentadoras de antígenos, células endoteliais ou células de órgãos-alvo para aGVHD⁽¹⁴⁰⁾. O gradiente de quimiocinas presente em órgãos-alvo leva a mais recrutamento de células efectoras, levando a lesão tissular e manifestação clínica de aGVHD^(141,142).

Um dos genes revelados no nosso estudo com função desconhecida no GVHD é o gene *TMEM176B* ou *LR8*, recentemente descrito em um modelo murino por similaridade de homologia como *TORID*⁽¹⁴³⁾. Este gene aparece preferencialmente expresso em células de linhagem mielóide, incluindo macrófagos e células dendríticas. Neste trabalho foi observado um aumento da expressão de *TORID* em células infiltrando enxerto cardíaco. Além disso, a super-expressão de *TORID* aumentou a expressão de moléculas de HLA classe II e de CD86 e diminuiu a produção de *IL-12p40*. Após ativação macrofágica induzida por lipopolissacáride, a expressão deste gene diminui consideravelmente, tanto em células dendríticas de camundongo quanto em células dendríticas derivadas de monócitos humanos. Estes dados sugerem que este gene possa estar envolvido no controle de captura antigênica e da diferenciação de células apresentadoras de antígenos imaturas. O fato de estar expresso principalmente em células dendríticas imaturas sugere que *TORID* promova uma apresentação tolerogênica de antígenos, não induzindo resposta imune.

Em nossa lista, o gene *TMEM176B* ou *LR8*, aparece com aumento de expressão no grupo de pacientes que evoluíram com aGVHD. Embora ainda não descrito sua função em humanos, podemos supor que o aumento da expressão deste gene, neste grupo de pacientes, possa ser secundária a uma tentativa prematura de freiar o processo

aloimune, através do controle de maturação de células dendríticas derivadas de medula óssea, na tentativa de induzir ou manter a tolerância imunológica.

Entre os genes de nossa lista com diminuição de expressão antes de aGVHD, a interleucina 1 é, provavelmente, a mais amplamente estudada. A família *IL-1* contém três produtos gênicos relacionados: *IL-1 α* , *IL-1 β* e *IL-1Ra* (receptor antagonista). *IL-1 α* e *IL-1 β* são protótipos de mediadores pró-inflamatórios que agem através de seu receptor (*IL-1R*), que em associação com proteínas acessórias de *IL-1R*, sinalizam através da proteína adaptadora *MyD88*⁽¹⁴⁴⁾. *IL-1 β* é expressa predominantemente em monócitos e macrófagos, enquanto *IL-1 α* é amplamente expressa em vários tipos celulares. O gene *RIPK2* codifica uma serino/treonina kinase^(145,146) que é essencial para a ativação de *NF- κ B* através do receptor de fator de necrose tumoral(*TNF*)⁽¹⁴⁷⁾, além de ser requerido para produção de *IL-2*, produção de *TCR* (receptor de células T) e diferenciação de células T auxiliares, sendo necessário para a diferenciação de células T e eficiente sinalização de *TCR*⁽¹⁴⁸⁾.

Nós observamos uma diminuição de expressão *IL-1* alfa, *IL-1* Beta e *RIPK2* em pacientes que evoluíram com aGVHD. A associação de aumento de *IL-1* e aGVHD é um tanto controversa. Células mononucleares contendo *IL-1*, *IL-1 β* e *TNF α* têm sido associadas com aGVHD, embora a concentração plasmática das mesmas citocinas tenha menor correlação com a doença⁽¹⁴⁹⁻¹⁵¹⁾. Em outro estudo, não foi observada diferença no aumento de expressão de mRNA de *IL-1*, *IL-2*, *IL-6* e *IFN γ* obtido por biópsia de pele de pacientes com aGVHD e comparado com pacientes sem aGVHD⁽¹⁵²⁾. Um estudo alemão observou que níveis aumentados de *IL-10*, mas não *IL-1* foi correlacionado com aGVHD⁽¹⁵³⁾. Estes dados conflitantes aqui discutidos podem ser explicados por vários motivos: períodos diferentes de coleta de amostras, fonte de amostras diferentes (sangue, biópsia de pele, fígado ou intestino), subtipos celulares utilizados, detecção de moléculas diferentes (mRNA, proteína) e estudos com pequeno número de pacientes com doenças e tratamentos heterogêneos.

Embora estes dados sejam conflitantes, é atribuído a *IL-1* um papel importante durante a fase efetora de aGVHD⁽⁵⁾. Secreção de *IL-1* parece ocorrer predominantemente

durante a fase efetora de aGVHD no baço e pele, órgãos-alvo da doença⁽¹⁵⁴⁾. Embora a administração de *IL1-Ra* a receptores reduza a mortalidade de aGVHD em modelos animais^(155,156), um recente trabalho randomizado em humanos não demonstrou qualquer benefício contra aGVHD⁽¹⁵⁷⁾.

As categorias do GO e a análise de redes de interação demonstrou o envolvimento de *IL-1* e *RIPK2* na ativação de *NF-κB* (Figura 11). *NF-κB* é uma fator de transcrição central que regula o desenvolvimento de linhagens celulares, diferenciação, sobrevivência e progressão de ciclo celular em muitos tipo celulares, incluindo células hematopoéticas⁽¹⁵⁸⁾. Por exemplo, a ativação de *NF-κB* tem sido associado a efeitos anti-apoptóticos em granulócitos⁽¹⁵⁹⁾. Outros genes na rede de comunicação bioquímica que nós observamos (Fig.11) induzidos por *IL-1* parecem também manter a diferenciação e sobrevivência celular. De fato, a indução de *PBEF1* (*NAMPT* na Fig. 11) por *IL-1β* em neutrófilos tem sido associado à postergação da apoptose destas células in vitro e em vivo⁽¹⁶⁰⁾. *KLF5* é um importante fator de transcrição que promove crescimento e proliferação do epitélio intestinal, além de ser encontrado em linfócitos T em desenvolvimento^(161,162). Receptor de hormônio nuclear *NR4A3* foi encontrado em células T imaturas e implicado na sobrevivência de múltiplos tipos celulares⁽¹⁶³⁾. Então, a ausência de expressão destes genes e a deficiência de ativação de *NF-κB* pode indicar um anormal desenvolvimento e ativação de células hematopoéticas em pacientes que desenvolverão aGVHD. Interessantemente, em um modelo experimental de aGVHD, o bloqueio de *NF-κB* em determinado período resultou em significativa aceleração da morbidade dependente de aGVHD⁽¹⁶⁴⁾. Nós também observamos que parte destes genes estão normalmente expresso em sangue de indivíduos saudáveis (Figura 13 cluster D). Este resultado indica que uma população de células “normais” deve estar sub-representada no sangue de pacientes que desenvolverão aGVHD, o que pode ser devido a aumento relativo nas células que tem baixa expressão desta assinatura.

Baseado neste dados, a expressão gênica do *cluster D* e a semelhança com expressão gênica de mesmos tipos celulares nos leva a acreditar na presença de diferentes subpopulações nestes dois grupos de pacientes, fato este somente elucidado

com caracterização imunofenotípica. De modo indireto, indica que subtipos celulares de células dendríticas BDCA4+, células CD34+ provenientes de medula óssea e monócitos elicitam respostas imune diferentes, gerando lesão tissular no grupo com aGVHD ou agindo de forma imunossupressora no grupo sem aGVHD. Melhor caracterização destes subtipos celulares representam uma potencial arma preditiva de ocorrência de aGVHD.

Em resumo, este trabalho descreveu a assinatura molecular preditiva de aGVHD, capaz de diagnosticar precocemente, em 80% da população estudada, a ocorrência de aGVHD. Este tipo de método diagnóstico precoce poderá permitir no futuro, intervenções apropriadas para cada situação clínica, oferecendo ao paciente um medicina personalizada com maior chance de sucesso. Em adição, estudamos genes diferencialmente expressos que foram agrupados em diferentes categorias funcionais. Nossos dados apontam para a potencial participação precoce das alarminas e da via IL-1/NF- κ B na resposta aloimune de aGVHD. Considerando os resultados da análise de enriquecimento funcional e a expressão de genes em diferentes populações celulares, nós sugerimos que eventos precoces envolvendo múltiplas populações de células precursoras possam predefinir a interação futura entre o enxerto em desenvolvimento e o paciente. Então, análise de expressão gênica de diferentes populações separadas em períodos precoces após TCTH serão necessárias para melhor delinear quais tipos celulares têm papel de destaque no desenvolvimento da aGVHD.

6. Conclusões

Nesse estudo de perfil de expressão gênica durante a enxertia medular, em pacientes submetidos a TCTH alogênico HLA-idêntico, nós observamos :

- ✓ um preditor molecular de aGVHD composto por 233 genes com acurácia de 75%.
- ✓ a expressão aumentada de alarminas, em pacientes que evoluíram com aGVHD, que podem representar “sinais de perigo” antecedendo a emergência de aGVHD.
- ✓ a expressão diminuída da IL1 e outros genes envolvidos na via do NF- κ B, em pacientes que evoluíram com aGVHD
- ✓ eventos precoces envolvendo múltiplas populações de células precursoras que podem predefinir a interação futura entre o enxerto em desenvolvimento e o paciente.

7. Anexos

Anexo 1: Pre-transplant high TREC levels indicate good prognosis after hematopoietic stem cell transplantation

Pre-transplant high TREC levels indicate good prognosis after hematopoietic stem cell transplantation

Adriano de Moraes Arantes^{1,2}, Kharen Kawemura¹, Adriana Seber³, José Salvador Rodrigues de Oliveira^{2,4}, Maria Gerbase DeLima¹, Andrey Morgun¹, Natasha Shulzhenko¹

Institutions: 1. Laboratory of Immunogenetics , Pediatrics Division, Universidade Federal de São Paulo; 2. Bone Marrow Transplant Unit, Universidade Federal de São Paulo ; 3. Pediatric Bone Marrow Transplant Center, Pediatric Oncology Institute - GRAACC, Universidade Federal de São Paulo; 4. Santa Marcelina Hospital

Adress: Rua Loefgreen n°1035, São Paulo,

ZIP-Code: 04040-031; São Paulo- Brazil

Body word count: 2937

Abstract word count: 193

The study was supported by a research grant from the AFIP. The sponsor of the study had no role in study design, data collection, data analysis, interpretation or writing of the report.

Correspondence address to: anemorgun@hotmail.com

Abstract**Word count: 193**

Background: Thymus-dependent T-cell reconstitution plays a role in immune recovery after stem cell transplantation (HSCT). High pre-HSCT thymic function has been associated with higher survival, lower incidence of acute and chronic graft versus host disease (GVHD) and lower incidence of infections. The aim of this study was to analyze the relationship between pre-HSCT peripheral blood levels of T cell receptor excision circles (TREC) and post-HSCT clinical events in recipients of HLA-identical hematopoietic stem cell transplants. **Method.** Delta deletion sjTRECs formed by the δ REC- ψ J α rearrangement were quantified by real time PCR in peripheral blood lymphocytes of 62 HSCT recipients. **Results.** Univariate analysis revealed association between low TREC levels and higher incidence of grade II-IV acute GVHD (p:.026), bacterial infection (p:.005) and cytomegalovirus infection (p:.033), whereas high TREC levels were associated with higher overall survival (p:0.028). In the multivariate analysis, low pre-HSCT TREC levels remained independently associated with lower survival (p: .032; RR 2.6), occurrence of grade II-IV acute GVHD (p:.031; RR: 2.5), bacterial infection (p:0.006, RR: 6.6) and cytomegalovirus infection (p:0.039; RR:2.8). **Conclusion:** Our results corroborate the concept that pre-HSCT recipient's thymic function is an important predictor of risk for acute grade II-IV GVHD and infection.

Keywords: T cell receptor excision circle, thymic function, stem cell transplantation

Introduction

Allogeneic hematopoietic stem cell transplantation (HSCT) is an effective treatment for hematological disorders; it is however burdened by high morbidity and mortality, mainly related to opportunistic infections and the occurrence of acute and chronic graft versus host disease. The immediate post transplant period is followed by a severe and often prolonged immune deficiency that results in persistent susceptibility to infection [Marks et al, 1993; Hansen et al, 1998; Hongeng et al, 1997; Kernan et al 1993]. Even more prolonged immune deficiency arises from inadequacy in effective CD4+ T cell and B cell reconstitution [Small et al, 1997; Small et al, 1999; Storek et al, 2000; Storek et al, 1997]. Regeneration of adequate T-cell number and repertoire diversity are key elements in the recovery of immune competence. In the absence of such recovery, susceptibility to infection is increased [Talvensaari et al, 2002].

Post transplant T cells are derived from both mature T cells present in the donor graft, and T cell that develop *de novo* in the recipient from transplanted donor stem cell. It is likely that the latter pathway of differentiation leads to long-term immune reconstitution [Weinberg et al, 2001; Rocha et al, 1989; Douek et al, 2000]. It has been suggested that the reconstitution and maintenance of effective T cell immunity after HSCT is dependent on education of T cell precursor in the thymus, as shown by the correlation between increased thymic output and an increment in naïve T cell number and broader TCR repertoires after transplantation [Dumont et al, 1998; Mackall et al, 1995; Roux et al, 2000]. Factors that inhibit thymic function may include age, chronic graft versus host disease (GVHD) and direct thymic damage from chemoradiotherapy [Douek et al, 1998; Nailor et al, 2005; Weinberg et al, 2001]. Age is a recognized predictor of low thymic function output and slow naïve T cell recovery after HSCT [Falen et al, 2003].

The production of a functional T cell receptor requires deletion of the TCR δ gene, which is positioned within the TCR α - locus [Livak et al, 1997]. The deleted TCR δ -genes remain present as extrachromosomal circular products named T-cell receptor excision circles - TRECs. These products are stable [Livak et al, 1996], are not duplicated during mitosis, and are therefore diluted out with each cellular division [Takeshita et al, 1989]. Quantification of TREC (T-cell receptor excision circles) levels after HSCT by real time PCR represents a sensitive measurement of thymopoietic capacity and has been used as a way to assess thymic function independently of cell surface phenotype [Harris et al, 2005, Hazenberg et al, 2000]. Recently it has been reported that TREC levels correlated to post-transplant outcome [Clave et al, 2005; Chen et al, 2005]. Herein we tested whether pre-transplant thymic function, assessed by the quantification of TRECs in the peripheral blood, may predict a worse prognosis in the setting of HLA-identical HSCT with a large prevalence of high risk disease. Furthermore, we also assessed TREC levels before and during acute GVHD in order to better understand thymic function at this period.

Materials and Methods

Patients, donors and transplantation characteristics

Over a consecutive 54-month period, a total of 62 allogeneic stem cell transplant recipients from Federal University of São Paulo (Bone Marrow Transplant Unit of Hospital Sao Paulo and Pediatric Bone Marrow Transplant Unit of the Pediatric Institute of Oncology) and Santa Marcelina Hospital, Brazil, between May 2003 and November 2007, were included. The group comprised children and adults who received a myeloablative preparative regimen and unmanipulated bone marrow or peripheral blood stem cell graft from a HLA-matched related

donor. Patients who received a non-myeloablative conditioning were excluded. Risk of disease was classified based in International Bone Marrow Transplant Registre (IBMTR) criteria (www.cibmtr.org). Patients' characteristics, treatment and complications after transplantation are summarized in Table 1. This study was approved by the Institutional Medical Ethics Committee and all patients or guardians gave informed consent before their enrollment.

Graft –versus-host disease prophylaxis, supportive care and end points

GVHD prophylaxis consisted predominantly of cyclosporine or cyclosporine and short course methotrexate. Cyclosporine levels were monitored weekly and kept at therapeutics range of 150-300 ng/dl during the first 60 days, and then the drug dose was continuously tapered through the next 180 days, until a complete suspension, depending on disease status at time of transplantation and presence or absence of GVHD. Antibacterial, antifungal, and antiviral prophylaxis consisted of trimethoprim-sulfamethoxazol for *Pneumocistis jiroveci*, acyclovir or valcyclovir for Herpes simplex. Surveillance cytomegalovirus (CMV) antigenemia testing was performed for all patients. Diagnosis and clinical grading of acute and chronic GVHD were performed according to established criteria [Glucksberg et al, 1974; Sullivan et al, 1991; Przepiorcka et al, 1995]. Bacterial, cytomegalovirus and fungal infection, sepsis and septic shock were defined according to previously established criteria [Ljungman et al, 1993; Society of Critical Care Medicine Consensus Conference Committee, 1992; Lossos et al, 1995].

Peripheral blood collection and time points

Buffy coat samples were obtained and frozen before the conditioning regimen for the 62 patients. 31 patients developed acute GVHD, and we were able to collect samples within 2 weeks before acute GVHD and ongoing GVHD in 16 from these patients. A control group was composed of

27 healthy donors from whom peripheral blood or bone marrow was collected before G-CSF stimulation

Clinical data

Between May 2003 and November 2007, 62 patients received a myeloablative conditioning regimen followed by allogeneic HCT at our center (table 1). During the follow up period of 18 months, among the total group of 62 patients, there were 14 severe bacterial infections, 22 CMV infections, 6 relapses and 3 fungal infections. Acute GVHD (aGVHD) occurred in 31 patients, and the incidence of grade I, II and III-IV was 13%, 18% and 19%, respectively. Among the 52 (83.8%) who survived for at least 100 days after transplantation, 34 (65%) developed chronic GVHD (cGVHD). As the number of relapses and fungal infection were low in our cohort, we focused the analysis on overall survival and occurrence of GVHD and severe bacterial infections.

Relative quantification of sj TREC by real-time –PCR

To perform TREC assay, total DNA was extracted from buffy-coat by DTAB/CTAB method (*dodecyl trimethyl-ammonium bromide/cetyl trimethyl-ammonium bromide*)-[Gustincich e al, 1991] and DNA concentrations were adjusted to 100ng/ul. One μ l of DNA was used in all reactions. Delta deletion sjTRECs formed by δ REC- ψ J α rearrangement were amplified and quantified using SYBR-Green PCR Master Mix and ABI 7000 instrument (Applied Biosystems, Foster City, CA). Standard curve was generated as previously described (PMID: 11295466). Primers were designed based on the gene sequence deposited in the GenBank (<http://www.ncbi.nlm.nih.gov>). Primers for sjTRECs were: forward primer

(5'CATCCCTTTCAACCATGC'3) and reverse primer (5'CGTGCCTAAACCCTGCAGC'3), which produced an approximately 102 base pair product detected. To improve TREC detection, two steps amplification was performed. The first PCR conditions were: 50°C for 2 minutes, 95°C for 10 minutes followed by 10 cycles of 95°C for 15 seconds and 63°C for 1 minute. After, 1ul was utilized in second step amplification. The second PCR conditions were: 95°C for 10 minutes followed by 40 cycles of 95°C for 15 seconds and 63°C for 1 minute. TREC values obtained by RT-PCR were normalized by proportion of lymphocytes in the blood cell count. This count was performed in the same peripheral blood sample that was utilized for DNA isolation. All DNA samples were amplified in triplicate and universal PCR precautions were taken to avoid amplicon contaminations along with negative control reactions.

Statistical analysis

Receiver operating characteristic (ROC) curve was used to obtain the most discriminating overall survival cut-off for TREC values. Correlation between Gaussian distribution variables was calculated with Spearman correlation. Survival analysis was calculated using Kaplan Meier plot and log rank test. Cumulative incidence analysis was calculated using death as a competing event for infections, acute or chronic GVHD. Cox proportional hazard model analysis was used to identify independent risk factors for death, infection, and acute or chronic GVHD, reviewed in the context of hematopoietic cell transplantation [Klein et al, 2001]. The variables included in the model were those that were marginally significant, p -value ≤ 0.10 , in the univariate analyses. Non-parametric (Wilcoxon and Mann-Whitney) tests were utilized to compare TREC levels between patients with high and low disease risk and before and during ongoing acute GVHD. All p values were two tailed and a value <0.05 was considered to indicate statistical significance. The analysis was done using the SPSS 10.0 Statistical Software (Chicago, Illinois, 1992).

Results

Correlation between TREC levels and age

Because in this study we used a modified assay for TREC detection, we wanted to confirm that it gives similar results to previous studies [Morgun et al, 2004, Svaldi et al, 2003]. To this end, we analyzed relation between TREC levels and age and, as expected, observed a highly significant negative correlation in healthy control individuals as well as in patients before transplantation ($r=-0.46$ and $r=-0.70$, respectively; $p\leq 0.0001$).

TREC levels and high/low risk disease

Although the pre-transplant TREC levels have been previously associated with post-transplant prognosis [Clave et al, 2005], we sought to repeat this analysis in our population of patients because it had a much bigger proportion of malignant high risk diseases (27% vs 1% in Clave et al, 2005). Indeed, we observed higher levels of TRECs in patients with no high risk disease (nonmalignant, low and intermediate risk) comparing to high risk disease patients in both adult and children populations ($p<0.001$, Figures 5A and 5B), thus making critical the inclusion of this factor in the further analysis.

TREC levels and survival

Using receiver operating characteristic (ROC) curve, we found a cut-off value of 134 TREC/ μg of lymphocyte DNA that discriminated between patients that survived or not at least 6 months after transplantation (65% of sensitivity and 60% of specificity). Based on this cut-off, patients were divided into “high” and “low” TREC groups. We observed higher survival in “high” TREC group ($p=0.028$) (Figure1) with almost twice more survivors in “high” TREC group vs “low” TREC group at 18 months after transplantation (65% vs. 37%, respectively). Applying a

multivariate analysis including TREC level, cell source (peripheral blood), recipient age (>25 years), ABO incompatibility, sex mismatch, and high risk disease, we found that worse survival was independently associated ($p<0.05$) to low TREC values before transplantation (RR: 2.6) and high risk disease (RR: 2.2) (Table 2).

We also observed a higher incidence of grades II-IV aGVHD (Figure 2) among low TREC patients ($p=0.02$), with incidence of aGVHD grade II-IV in high and low TREC groups of 25.7% and 57.8%, respectively. Multivariate analysis showed that low TREC levels (RR 2.5, $p<0.03$) and the antecedent of high risk disease (RR: 5.2 $p<0.025$) were independently associated with increased incidence of aGVHD grades II-IV (Table 2).

Although we did not find association between TREC levels and cGVHD, the expected associations of cGVHD with age [Ringden et al, 1985; Atkinson et al, 1990, Kondo et al, 2001] and source of cells for transplant [Mohty et al, 2002; Anderson et al, 2003; Cutler et al, 2001] were detected (Table 2).

TREC levels and infections

Next we performed the same analysis for bacterial and cytomegalovirus infection. There was higher incidence of severe bacterial infection in the low TREC (42%) compared to the high TREC (8.5%) group ($p=0.005$) (Figure 3). Considering other variables, only low TREC level and peripheral blood as a source of transplant independently associated to bacterial infections (RRs: 6.6 and 4.2, respectively) (Table 2). CMV infection was also more frequently observed in the low TREC patients (63% vs. 22%, $p<0.03$) (Figure 4). Furthermore, low TREC values before transplantation and high risk disease had independent impact on the incidence of CMV infection (Table 2).

Post-transplant TREC levels in relation to acute GVHD

Having the result that pre-transplant thymus function was associated with aGVHD, we wondered if thymus output after transplantation (i.e. TREC levels) would also be related to aGVHD. For this purpose we assessed TRECs levels in blood samples collected before (pre-aGVHD) and during aGVHD from 16 patients. The time period between pre-aGVHD and aGVHD samples was 11.5 ± 8 days. Comparing TREC levels between pre-aGVHD and aGVHD samples, we observed increased levels during aGVHD (147 vs. 480 TRECs / μ g of lymphocyte DNA, $p=0.015$) (Figure 2). One could suggest that this difference could be explained by gradual increase in TREC levels after transplantation. In which case, the larger time distance between pre-aGVHD and GVHD samples, the larger would be the increment in TREC levels. The analysis, however, showed no correlation between distance in time and TREC levels (data not shown). In addition, there was no correlation between number of days after transplantation (17 to 49 days) and TREC levels in this time period. This indicates that we have observed increase in TREC levels associated with aGVHD independently on overall increase in TREC levels after transplantation which occurs more evidently after the third post-transplant month [Hochberg et al, 2001; Chen et al, 2005].

Discussion

It has been postulated that a good immune reconstitution after HSCT means a better prognosis for the patients under stem cell transplantation [Parkman et al, 2006; Kim et al, 2006]. The evaluation of immune reconstitution assessed by TREC quantification has been a matter of research, it is now possible to determine the impact of this variable in the clinical settings. Here, we analyzed the impact of pre-transplant thymic function in patients undergoing a HSCT. It has

been recently reported that TREC levels correlated to post-transplant outcome [Clave et al, 2005; Chen et al, 2005]. However, it may be suggested that TREC levels represent a risk factor for post-transplant outcomes secondary to high/low risk disease patient status, especially considering a strong relation between disease risk and TREC levels (Figure 5A and 5B). Differently from previously published work [Clave et al, 2005], we had analyzed a cohort of patients with much higher proportion of malignant high risk diseases that allowed us to address the above mentioned concern. In fact, we observed similar results to the previously published work showing TREC levels as a surrogate marker for complication after HSCT independently on the disease risk.

Interestingly, we found that low TREC levels were associated with higher probability of bacterial infection and aGVHD. This suggests that TREC levels are not just associated with high or low T cell response after transplantation but rather reflect an overall response to damage as well as potential capacity to establish immune homeostasis after transplantation. Indeed, in the setting of high risk disease, patients were treated prior to transplantation with multiples cycles of chemotherapy and/or radiotherapy, potential contributors to thymic injury. Cortical and medullary thymic epithelial cells are dramatically altered after radiotherapy and chemotherapy [Adkins et al, 1988, Mackall et al, 1995] and damage in this thymic stromal compartment affects functions as attraction of hematopoietic precursors to the thymus and functional maturation of newly generated T cell precluding normal thymocyte development [Hollander et al, 1995] and probably TRECs generation.

Diminished levels of TRECs has been related in patients with ongoing chronic GVHD or with an previous history of acute or chronic GVHD [Weinberg et al, 2001; Talvensaari et al, 2002], although others studies have failed to demonstrate a direct association of TRECs levels with acute or chronic GVHD [Storek et al, 2001; Hochberg et al 2001]. In a mouse model of aGVHD, TREC levels were analyzed in spleen and thymus but not in bloods making the direct comparison of these experiments in mice with our results complicate [Krenger et al, 2004].

TREC levels during acute GVHD have not been reported in patients. With this purpose, we performed this analysis and observed an increase of TRECs levels during ongoing acute GVHD in relation to two weeks pre-aGVHD. Interestingly, herein observed increase in TREC levels during acute GVHD is very similar to our previous observation of increased TREC levels during acute cardiac rejection [Morgun et al, 2004] suggesting that similar events may be going on in these two pathological processes in blood. The explanation of these results, on one hand, could be an increased output of TREC positive naive T cells into circulation, stimulated by the ongoing immune response [Chen et al, 2005]. On the other hand, this result might indicate the exit of activated/effector cells from the circulation into target tissues and a consequent increase in naïve T cell pool in the blood. Having in mind the limitations of our analysis (small number of patients and indirect comparison), we think that additional research is necessary in order to draw definitive conclusions about changes in post-transplant TREC levels in relation to aGVHD.

In this work we could demonstrate a relatively easy way of normalizing TREC values determined by RT-PCR, without the concomitant need for flow cytometry to estimate numbers of CD3+ cells. And it was interesting to notice that a cut-off value established by ROC-curve analysis, which could discriminate between good or bad prognosis, was similar to the value found at 60th percentile, after TREC values distribution in percentiles from published works [Clave et al, 2005; Chen et al, 2005]. Using our cut-off, we could observe very similar to the previously published results in relation to overall survival, acute GVHD, bacterial and CMV infections [Clave E et al 2005]. It was a interesting idea that a cut-off value could be determined by distributing values found in a group patients, independent of the methods used for its determination. Finally, we have shown a disease risk as one the most important sources of TREC levels variability in patients before HSCT. Further studies should consider this factor along with other factors such as age in order to establish how TREC levels can be used in diagnostic practice.

In conclusion, our results further support the idea that the thymic competence before transplantation is a critical factor for a good prognosis after hematopoietic stem cell transplantation.

Table 1. Patients' characteristics.

Characteristics	N:62
Recipient age, median (range)	26(2-58)
Age over 25 y	31
Diagnosis, n°	
Acute Lymphoblastic Leukemia	12
Acute Myeloid Leukemia	16
Myelodysplastic Syndrome	7
Aplastic Anemia	6
Hodgkin disease	2
Chronic Myeloid Leukemia	9
Non Hodgkin Lymphoma	10
Risk Disease*, n°	
Nonmalignant	
Malignant low risk	6
Malignant intermediate risk	25
Malignant high risk	14
	17
GVHD prophylaxys, n°	
CSA + MTX	42
CSA	13
Others	7
Conditioning Regimen, n°	
Bussulphan based	39
TBI based	13
Others	10
Engraftment in days , median (range)	14(10-25)
Median cells infused (range)	
TNC/Kg x10 ⁸	7.0(2.3-11)
CD34/Kg x10 ⁶	4.9(2.1-12)
Cells Source	
Bone Marrow	30
Peripheral blood stem cell	32
Major ABO incompatibility	13
Acute GVHD	
Absent	31
Grade I	8
Grade II-IV	23
Chronic GVHD	
Absent	18
Localized	11
Extensive	23
Cytomegalovirus reactivation	22
Severe Bacterial Infection	14
Fungal infection	3
Relapse	6
Follow up in months , median (range)	18(0.5-60)

CSA-cyclosporine, MTX-methotrexate, TBI-total body irradiation, TNC – total nucleated cell, GVHD – graft versus host disease.

Table 2. Cox Model Analysis

	Univariate p	RR	Multivariate CI	p
Overall Survival				
Low TREC	0.028	2.6	1.1-6.3	0.032
Age more than 25y	0.47	----	----	0.39
F→M vs others	0.64	----	----	----
Peripheral blood vs bone marrow	0.13	----	----	----
High risk disease vs others	0.05	2.2	1.1-4.7	0.044
ABO Incompatibility	0.26	----	----	----
Grade II-IV acute GVHD				
Low TREC	0.026	2.5	1.3-5.9	0.031
Age more than 25y	0.036	----	----	0.15
M→H vs others	0.035	----	----	0.94
Peripheral blood vs bone marrow	0.40	----	----	----
High risk disease vs others	0.02	5.2	1.2-5.9	0.025
ABO Incompatibility	0.26	----	----	----
Extensive chronic GVHD				
Low TREC	0.48	----	----	0.54
Age more than 25y	0.008	----	----	0.11
M→H vs others	0.42	----	----	0.14
Peripheral blood vs bone marrow	0.005	----	----	----
High risk disease vs others	0.51	----	----	----
ABO Incompatibility	0.19	----	----	----
Bacterial Infection				
Low TREC	0.005	6.6	1.7-25	0.006
Age more than 25y	0.090	----	----	0.95
M→H vs others	0.69	----	----	----
Peripheral blood vs bone marrow	0.037	4.2	1.2-15.6	0.028
High risk disease vs others	0.36	----	----	----
ABO Incompatibility	0.43	----	----	----
Cytomegalovirus infection				
Low TREC	0.033	2.8	1.4-5.7	0.039
Age more than 25y	0.15	----	----	0.59
M→H vs others	0.65	----	----	----
Peripheral blood vs bone marrow	0.93	----	----	----
High risk disease vs others	0.16	----	----	----
ABO Incompatibility	0.23	----	----	----

(F→M) denotes female donor into male recipient, RR indicates relative ratio
95% CI, 95% confidence interval

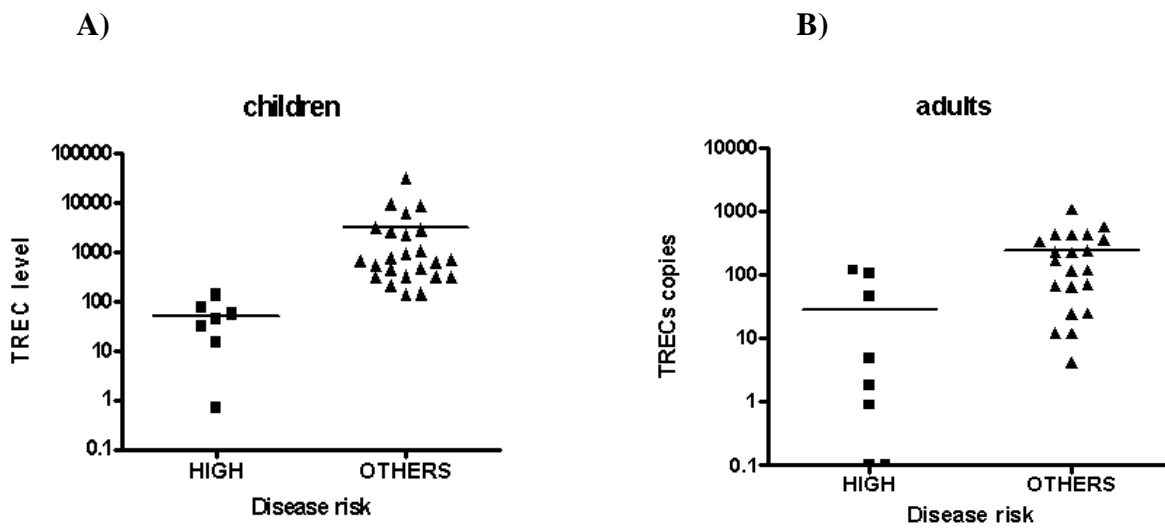


Figure 1A: Comparison of TRECs levels between high risk disease (N:7) and others (nonmalignant, low and intermediated risk disease ; N: 24) in children before stem cell transplantation ($p < .001$). 1B) Comparison of TRECs levels between high risk disease (N:10) and others and others (nonmalignant, low and intermediated risk disease; N: 21) and in adults before stem cell transplantation ($p < .001$)

Overall Survival

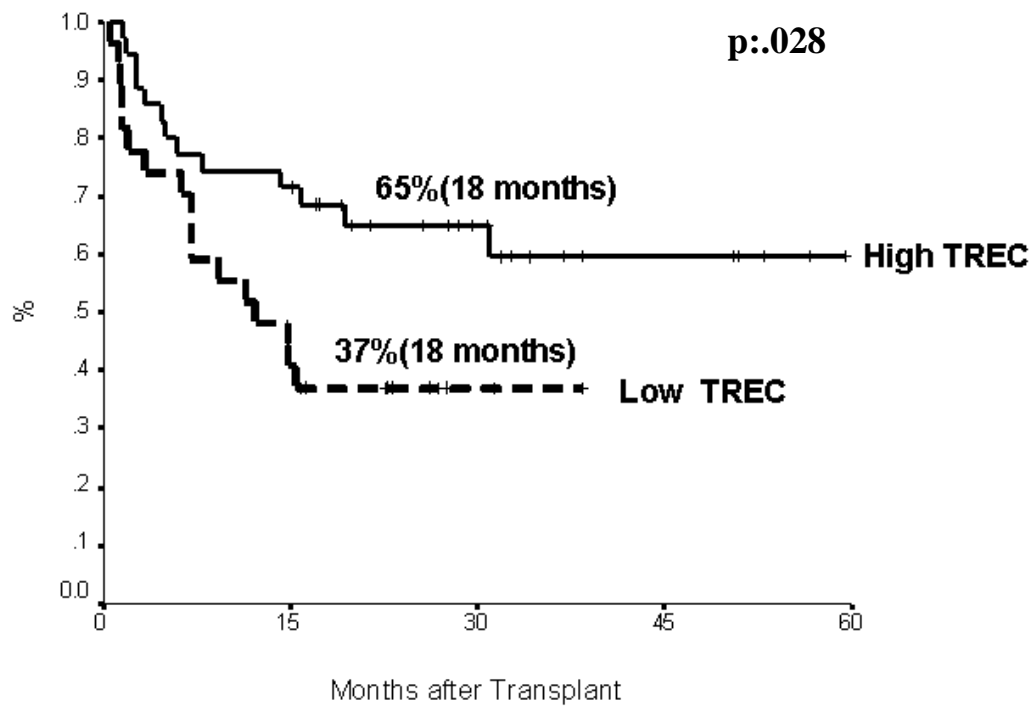


Figure 2. Overall Survival after stem cell transplantation stratified by TREC levels. Patients with high TREC (N:35) are shown as solid line while the patients with low TREC (N:27) is plotted as dashed line. Log rank P:028

Grade II-IV acute GVHD

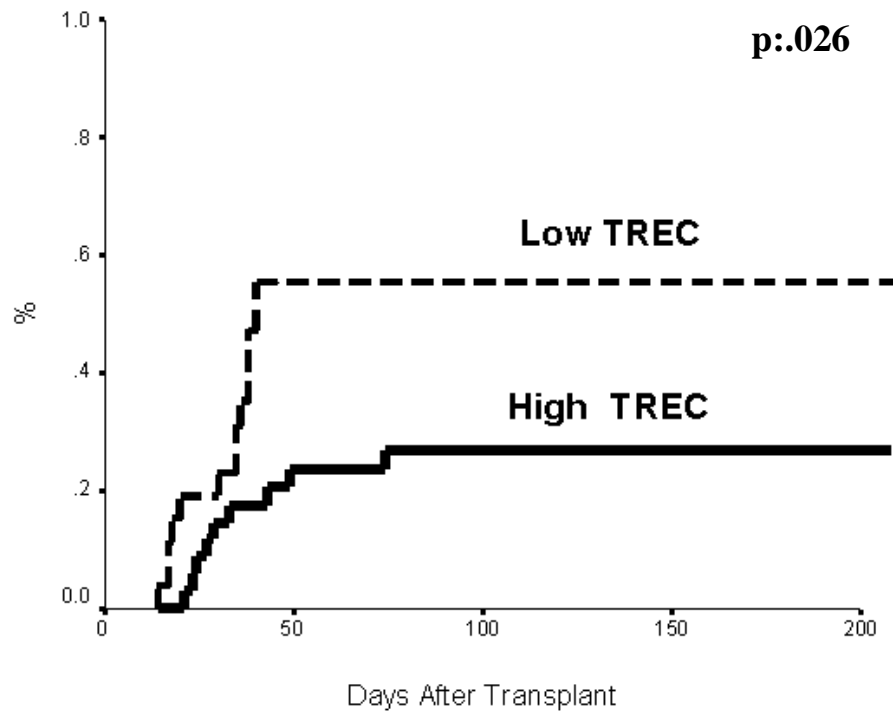


Figure 3: Cumulative incidence of acute GVHD in low level TREC (N:27) group compared with high level group (N:32). Patients with high TREC are shown as solid line while the patients with low TREC is plotted as dashed line. Log rank p:.026

Bacterial Infection

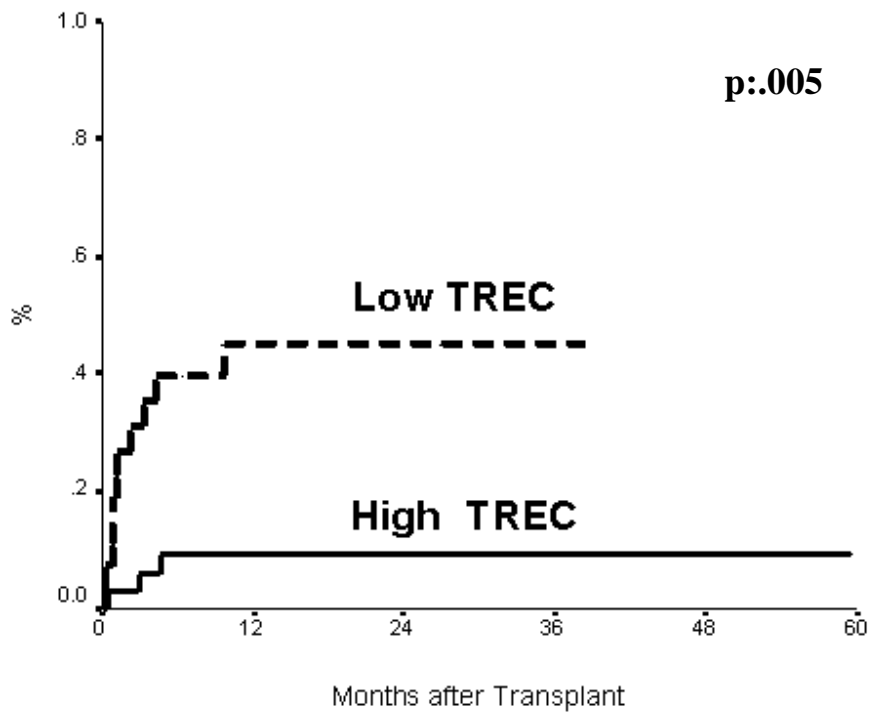


Figure 4: Cumulative incidence of bacterial infection in low level TREC (N:27) group compared with high level group (N:32). Patients with high TREC are shown as solid line while the patients with low TREC is plotted as dashed line Log rank p:.005

Cytomegalovirus Infection

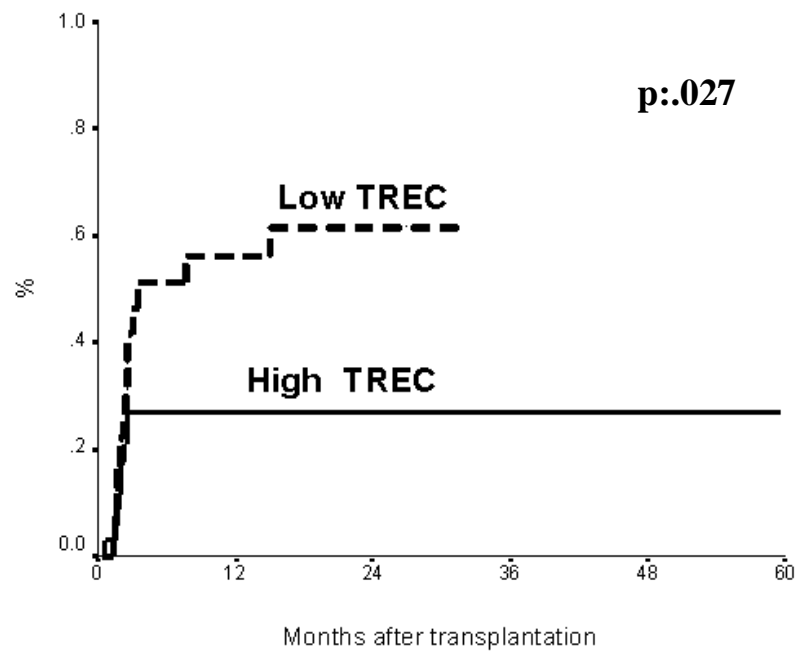


Figure 5: Cumulative incidence of cytomegalovirus infection in low level TREC (N:27) group compared with high level group (N:32). Patients with high TREC are shown as solid line while the patients with low TREC is plotted as dashed line Log rank p:.027

References

Adkins B, Gandour D, Strober S, Weissman I.(1988) Total lymphoid irradiation leads to transient depletion of the mouse thymic medulla and persistent abnormalities among medullary stromal cells. *J Immunol*, **140**,3373-9.

American Society of Chest Physicians/Society of Critical Care Medicine Consensus Conference Committee.(1992) Definitions of sepsis and organ failure and guidelines for the use of innovative therapies in sepsis. *Crit Care Med*, **20**:864-874.

Anderson D, DeFor T, Burns L, McGlave P, Miller J, Wagner J, Weisdorf D.(2003) A comparison of related donor peripheral blood and bone marrow transplants: importance of late-onset chronic graft-versus-host disease and infections. *Biol Blood Marrow Transplant*, **9**,52-9.

Atkinson K, Horowitz MM, Gale RP, van Bekkum DW, Gluckman E, Good RA, Jacobsen N, Kolb HJ, Rimm AA, Ringdén O.(1990) Risk factors for chronic graft-versus-host disease after HLA-identical sibling bone marrow transplantation. *Blood*, **75**,2459-64.

Chen X, Barfield R, Benaim E, Leung W, Knowles J, Lawrence D, Otto M, Shurtleff SA, Neale GA, Behm FG, Turner V, Handgretinger R.(2005) Prediction of T-cell reconstitution by assessment of T-cell receptor excision circle before allogeneic hematopoietic stem cell transplantation in pediatric patients. *Blood*,**105**,886-93.

Clave E, Rocha V, Talvensaaari K, Busson M, Douay C, Appert ML, Rabian C, Carmagnat M, Garnier F, Filion A, Socié G, Gluckman E, Charron D, Toubert A.(2005) Prognostic value of pretransplantation host thymic function in HLA-identical sibling hematopoietic stem cell transplantation. *Blood*, **105**,2608-13.

Cutler C, Giri S, Jeyapalan S, Paniagua D, Viswanathan A, Antin JH. Acute and chronic graft-versus-host disease after allogeneic peripheral-blood stem-cell and bone marrow transplantation: a meta-analysis. *J Clin Oncol*, **19**,3685-91.

Douek DC, McFarland RD, Keiser PH, Gage EA, Massey JM, Haynes BF, Polis MA, Haase AT, Feinberg MB, Sullivan JL, Jamieson BD, Zack JA, Picker LJ, Koup RA.(1998) Changes in thymic function with age and during the treatment of HIV infection. *Nature*, **396**:690-5.

Douek DC, Vescio RA, Betts MR, Brenchley JM, Hill BJ, Zhang L, Berenson JR, Collins RH, Koup RA.(2000) Assessment of thymic output in adults after haematopoietic stem-cell transplantation and prediction of T-cell reconstitution. *Lancet*, **355**,1875-81.

Dumont-Girard F, Roux E, van Lier RA, Hale G, Helg C, Chapuis B, Starobinski M, Roosnek E.(1998) Reconstitution of the T-cell compartment after bone marrow transplantation: restoration of the repertoire by thymic emigrants. *Blood*, **92**,4464-71.

Fallen PR, McGreavey L, Madrigal JA, Potter M, Ethell M, Prentice HG, Guimarães A, Travers PJ.(2003) Factors affecting reconstitution of the T cell compartment in allogeneic haematopoietic cell transplant recipients. *Bone Marrow Transplant*, **32**,1001-14.

Glucksberg H, Storb R, Fefer A, Buckner CD, Neiman PE, Clift RA, Lerner KG, Thomas ED.(1974) Clinical manifestations of graft-versus-host disease in human recipients of marrow from HL-A-matched sibling donors. *Transplantation*, 1974,**18**,295-304.

Gustincich S, Manfioletti G, Del Sal G, Schneider C, Carninci P.(1991) A fast method for high-quality genomic DNA extraction from whole human blood. *Biotechniques*,**11**,298-300.

Hansen JA, Gooley TA, Martin PJ, Appelbaum F, Chauncey TR, Clift RA, Petersdorf EW, Radich J, Sanders JE, Storb RF, Sullivan KM, Anasetti C.(1998) Bone marrow transplants from unrelated donors for patients with chronic myeloid leukemia. *N Engl J Med*, **338**,962-8

Harris JM, Hazenberg MD, Poulin JF, Higuera-Alhino D, Schmidt D, Gotway M, McCune JM.(2005) Multiparameter evaluation of human thymic function: interpretations and caveats. *Clin Immunol*, **115**,138-46.

Hazenberg MD, Otto SA, Cohen Stuart JW, Verschuren MC, Borleffs JC, Boucher CA, Coutinho RA, Lange JM, Rinke de Wit TF, Tsegaye A, van Dongen JJ, Hamann D, de Boer RJ, Miedema F.(2000) Increased cell division but not thymic dysfunction rapidly affects the T-cell receptor excision circle content of the naive T cell population in HIV-1 infection. *Nat Med*, **6**,1036-42.

Hochberg EP, Chillemi AC, Wu CJ, Neuberger D, Canning C, Hartman K, Alyea EP, Soiffer RJ, Kalams SA, Ritz J.(2001) Quantitation of T-cell neogenesis in vivo after allogeneic bone marrow transplantation in adults. *Blood*, **98**,1116-21.

Holländer GA, Wang B, Nichogiannopoulou A, Platenburg PP, van Ewijk W, Burakoff SJ, Gutierrez-Ramos JC, Terhorst C.(1995) Developmental control point in induction of thymic cortex regulated by a subpopulation of prothymocytes. *Nature*, **373**,350-3.

Hongeng S Lancet 1997, Hongeng S, Krance RA, Bowman LC, Srivastava DK, Cunningham JM, Horwitz EM, Brenner MK, Heslop HE.(1997) Outcomes of transplantation with matched-sibling and unrelated-donor bone marrow in children with leukaemia. *Lancet*, **350**,767-71.

Kernan NA, Bartsch G, Ash RC, Beatty PG, Champlin R, Filipovich A, Gajewski J, Hansen JA, Henslee-Downey J, McCullough J.(1993). Analysis of 462 transplantations from unrelated donors facilitated by the National Marrow Donor Program. *N England J Med*, **328**,593-602.

Kim DH, Sohn SK, Won DI, Lee NY, Suh JS, Lee KB.(2006) Rapid helper T-cell recovery above $200 \times 10^6/l$ at 3 months correlates to successful transplant outcomes after allogeneic stem cell transplantation. *Bone Marrow Transplant*, **37**,1119-28.

Klein JP, Rizzo JD, Zhang MJ, Keiding N.(2001) Statistical methods for the analysis and presentation of the results of bone marrow transplants. Part I: unadjusted analysis. *Bone Marrow Transplant*, **28**,909-15.

- Kondo M, Kojima S, Horibe K, Kato K, Matsuyama T.(2001) Risk factors for chronic graft-versus-host disease after allogeneic stem cell transplantation in children. *Bone Marrow Transplant*, **27**,727-30.
- Krenger W, Schmidlin H, Cavadini G, Holländer GA.(2004) On the relevance of TCR rearrangement circles as molecular markers for thymic output during experimental graft-versus-host disease. *J Immunol*, **172**,7359-67.
- Livak F, Schatz DG.(1996) T-cell receptor alpha locus V(D)J recombination by-products are abundant in thymocytes and mature T cells. *Mol Cell Biol*, **16**, 609-18.
- Livák F, Schatz DG.(1997) Identification of V(D)J recombination coding end intermediates in normal thymocytes. *J Mol Biol*, **67**,1-9.
- Ljungman P, Griffiths P. Definitions of cytomegalovirus infection and disease. (1993) Proceedings of the Fourth International Cytomegalovirus Workshop: Multidisciplinary Approach to Understanding Cytomegalovirus Disease. Paris, France: *Elsevier Science Publishers BV*, 233-237.
- Lossos IS, Breuer R, Or R, Strauss N, Elishoov H, Naparstek E, Aker M, Nagler A, Moses AE, Shapiro M.(1995). Bacterial pneumonia in recipients of bone marrow transplantation. A five-year prospective study. *Transplantation*, **60**,672-8.
- Mackall CL, Fleisher TA, Brown MR, Andrich MP, Chen CC, Feuerstein IM, Horowitz ME, Magrath IT, Shad AT, Steinberg SM.(1995) Age, thymopoiesis, and CD4+ T-lymphocyte regeneration after intensive chemotherapy. *N Engl J Med*, **332**,143-9.
- Marks DI, Cullis JO, Ward KN, Lacey S, Syzdlo R, Hughes TP, Schwarer AP, Lutz E, Barrett AJ, Hows JM, Batchelor JR, Goldman JM.(1993) Allogeneic bone marrow transplantation for chronic myeloid leukemia using sibling and volunteer unrelated donors. A comparison of complications in the first 2 years. *Ann Intern Med*, **119**,207-14
- Mohty M, Kuentz M, Michallet M, Bourhis JH, Milpied N, Sutton L, Jouet JP, Attal M, Bordigoni P, Cahn JY, Boiron JM, Blaise D; Société Française de Greffe de Moelle et de Thérapie Cellulaire (SFGM-TC).(2002) Chronic graft-versus-host disease after allogeneic blood stem cell transplantation: long-term results of a randomized study. *Blood*,**100**,3128-34.
- Morgun A, Shulzhenko N, Socorro-Silva A, Diniz RV, Almeida DR, Gerbase-Delima M.(2004) T cell receptor excision circles (TRECs) in relation to acute cardiac allograft rejection. *J Clin Immunol*, **24**,612-6.
- Naylor K, Li G, Vallejo AN, Lee WW, Koetz K, Bryl E, Witkowski J, Fulbright J, Weyand CM, Goronzy JJ. (2005)The influence of age on T cell generation and TCR diversity. *J Immunol*, **174**,7446-52.
- Parkman R, Cohen G, Carter SL, Weinberg KI, Masinsin B, Guinan E, Kurtzberg J, Wagner JE, Kernan NA.(2006) Successful immune reconstitution decreases leukemic relapse and improves survival in recipients of unrelated cord blood transplantation. *Biol Blood Marrow Transplant*, **12**,919-27.

Przepiorka D, Weisdorf D, Martin P, Klingemann HG, Beatty P, Hows J, Thomas ED.(1995) 1994 Consensus Conference on Acute GVHD Grading. *Bone Marrow Transplant*,**15**,825-8.

Ringdén O, Paulin T, Lönnqvist B, Nilsson B.(1985) An analysis of factors predisposing to chronic graft-versus-host disease. *Exp Hematol*, **13**,1062-7.

Rocha B Eur J Immunol 1989, Rocha B, Dautigny N, Pereira P.(1989) Peripheral T lymphocytes: expansion potential and homeostatic regulation of pool sizes and CD4/CD8 ratios in vivo. *Eur J immunol*,**19**,905-11.

Roux E, Dumont-Girard F, Starobinski M, Siegrist CA, Helg C, Chapuis B, Roosnek E.(2000) Recovery of immune reactivity after T-cell-depleted bone marrow transplantation depends on thymic activity. *Blood*, **96**,2299-303.

Small TN, Avigan D, Dupont B, Smith K, Black P, Heller G, Polyak T, O'Reilly RJ.(1997) Immune reconstitution following T-cell depleted bone marrow transplantation: effect of age and posttransplant graft rejection prophylaxis.*Biol Blood Marrow Transplan*, **3**,65-75.

Small TN, Papadopoulos EB, Boulad F, Black P, Castro-Malaspina H, Childs BH, Collins N, Gillio A, George D, Jakubowski A, Heller G, Fazzari M, Kernan N, MacKinnon S, Szabolcs P, Young JW, O'Reilly RJ.(1999) Comparison of immune reconstitution after unrelated and related T-cell-depleted bone marrow transplantation: effect of patient age and donor leukocyte infusions. *Blood*, **93**,467-80.

Storek J, Espino G, Dawson MA, Storer B, Flowers ME, Maloney DG.(2000) Low B-cell and monocyte counts on day 80 are associated with high infection rates between days 100 and 365 after allogeneic marrow transplantation. *Blood*, **96**,3290-3.

Storek J, Gooley T, Witherspoon RP, Sullivan KM, Storb R.(1997) Infectious morbidity in long-term survivors of allogeneic marrow transplantation is associated with low CD4 T cell counts. *Am J Hematol*, **54**,131-8.

Storek J, Joseph A, Espino G, Dawson MA, Douek DC, Sullivan KM, Flowers ME, Martin P, Mathioudakis G, Nash RA, Storb R, Appelbaum FR, Maloney DG.(2001) Immunity of patients surviving 20 to 30 years after allogeneic or syngeneic bone marrow transplantation. *Blood*, **98**,3505-12.

Sullivan KM, Agura E, Anasetti C, Appelbaum F, Badger C, Bearman S, Erickson K, Flowers M, Hansen J, Loughran T.(1991) Chronic graft-versus-host disease and other late complications of bone marrow transplantation. *Semin Hematol*, **28**,250-9

Svaldi M, Lanthaler AJ, Dugas M, Lohse P, Pescosta N, Straka C, Mitterer M.(2003) T-cell receptor excision circles: a novel prognostic parameter for the outcome of transplantation in multiple myeloma patients. *Br J Haematol*, **122**,795-801.

Takeshita S, Toda M, Yamagishi H.(1989) Excision products of the T cell receptor gene support a progressive rearrangement model of the alpha/delta locus. *EMBO J*, **8**,3261-70.

Talvensaari K, Clave E, Douay C, Rabian C, Garderet L, Busson M, Garnier F, Douek D, Gluckman E, Charron D, Toubert A.(2002) A broad T-cell repertoire diversity and an efficient thymic function indicate a favorable long-term immune reconstitution after cord blood stem cell transplantation. *Blood*, **99**,1458-64.

Weinberg K, Blazar BR, Wagner JE, Agura E, Hill BJ, Smogorzewska M, Koup RA, Betts MR, Collins RH, Douek DC.(2001) Factors affecting thymic function after allogeneic hematopoietic stem cell transplantation. [*Blood*](#), **97**,1458-66.

Anexo 2. Características clínicas dos pacientes incluídos no estudo

Número Paciente	Array	Amostra	SR	SD	IR (anos)	RC	RC ablação	Profilaxia GVHD	Doença	Status ao TMO	Risco da Doença	Fonte de célula	Incompat ABO	Incompat Sexo (F→M)
1	85499	TMO292	M	F	51	BuCy	MA	CSA+MTX	LMA com displasia multilinhagem	Remissão parcial	alto	SP	ausente	Sim
2	85500	TMO297	M	M	58	BuCy	MA	CSA+MTX	LMC	Fase Crônica	baixo	MO	ausente	Não
3	85501	TMO044	F	F	39	BuCy	MA	CSA+MTX	LMC	Fase acelerada	Interm	MO	ausente	Não
4	85523	TMO628	M	M	21	TBI+Cy	MA	CSA+MTX	LLA pró B	1ª Remissão completa	baixo	SP	ausente	Não
5	85544	TMO020	F	M	53	BuCy	MA	CSA+MTX	LMC	Fase Crônica	baixo	MO	ausente	Não
6	85551	TMO266	F	M	23	BuCy	MA	CSA+MTX	LMA -minimamente diferenciada (FAB-M0)	Refratária	alto	MO	ausente	Não
7	85552	TMO456	M	M	33	BuCy	MA	CSA+MTX	SMD AREB II	----	alto	SP	ausente	Não
8	85553	TMO409	M	M	54	FluBu	MA	CSA+MTX	LMA sem maturação (FAB-M1)	1ª Remissão completa	baixo	SP	maior	Não
9	85554	TMO362	F	M	56	Mel140	RTR	CSA+MTX	MM IgG kappa IIIA	Refratário	alto	SP	ausente	Não
10	85555	TMO076	M	M	51	BuCy	MA	CSA+MTX	LMA com displasia multilinhagem	Refratária	alto	SP	ausente	Não
11	85556	TMO524	F	F	50	FluBu	MA	CSA+MTX	LNH T/ Micose Fungóide T3N0M0	Refratário	alto	SP	ausente	Não
12	85557	TMO148	F	M	46	buCy	MA	CSA+MTX	LMMC Agudizada	1ª Remissão completa	alto	SP	menor	Não
13	85558	TMO433	M	F	31	BuCy	MA	CSA+MTX	LMA com displasia multilinhagem	1ª Remissão completa	baixo	SP	menor	Sim
14	85559	TMO523	F	F	23	BuCy	MA	CSA+MTX	LBA Monocítica/T	1ª Remissão completa	alto	SP	maior	Não
15	86801	TMO254	F	M	33	BuCy	MA	CSA+MTX	SMD hipocelular	----	baixo	MO	ausente	Não
16	85560	TMO588	F	F	29	BuCy	MA	CSA+MTX	SMPC inclassificável		interm	MO	ausente	Não
17	85575	TMO525	M	M	17	TBI+Cy	MA	CSA+MTX	LLA T	2ª Remissão completa	interm	SP	ausente	Não
18	85572	TMO219	F	M	46	CyMel	RTR	CSA+MTX	Linfoma Folicular Grau II IVB	3ª recidiva	alto	SP	ausente	Não
19	85581	TMO554	M	M	32	TBI+Cy	MA	CSA+MTX	LMA -minimamente diferenciada (FAB-M0)	Remissão parcial	alto	SP	ausente	Não
20	85582	TMO440	F	M	55	FluBu	MA	CSA+MTX	LBA monocítica/T	1ª Remissão completa	alto	SP	menor	Não
21	85583	TMO243	M	F	57	CyMel	RTR	CSA+MTX	LNH Linfoplasmocitóide IVB	Refratário	alto	MO	ausente	Sim
22	85584	TMO457	M	F	25	CyMel	RTR	CSA+MTX	LHEN IVB	recidiva após autólogo	alto	SP	ausente	Sim
23	85591	TMO240	F	F	49	BuCy	MA	CSA+MTX	HPN	Fase hipocelular	interm	MO	ausente	Não

Número Paciente	Array	Amostra	Exertia medular(dias)	aGVHD - Graduação	aGVHD por sítio	Biópsia-aGVHD	Diagnóstico aGVHD (dias)	cGVHD extensão	Diagnóstico cGVHD (dias)	cGVHD tipo de início	Seg(dias)	Status	Causa do óbito
1	85499	TMO292	13	grau II	Fígado G1/ Pele G1	sim-pele	38	extenso	100	progressivo	277	morto	GVHD crônico + Sepse
2	85500	TMO297	18	grau II	Fígado G1/ Pele G2	não	43	ausência	----	-----	81	morto	GVHD agudo + Pnemonite por CMV
3	85501	TMO044	12	grau III	Pele G2 /fígadoG2 / TGI G2	Não	21	extenso	100	progressivo	1037	vivo	----
4	85523	TMO628	12	grau I	Pele G2	sim-pele	33	extenso	223	quiescente	631	vivo	----
5	85544	TMO020	19	ausência	----	----	----	extenso	132	de novo	594	morto	GVHD crônico
6	85551	TMO266	17	grau III	Fígado G2 /TGI G2	sim-intestino	49	extenso	159	quiescente	938	vivo	----
7	85552	TMO456	14	ausência	----	----	----	extenso	114	de novo	805	vivo	----
8	85553	TMO409	12	ausência	----	----	----	extenso	161	de novo	871	vivo	----
9	85554	TMO362	12	grau III	Pele G3/ Fígado G2 /TGIG2	Não	22	ausência	----	-----	189	morto	Pnemonite por CMV
10	85555	TMO076	14	grau III	TGI G2 /Fígado G2	Não	23	ausência	----	-----	38	morto	GVHD agudo + Sepse
11	85556	TMO524	11	ausência	----	----	----	extenso	83	de novo	723	vivo	----
12	85557	TMO148	16	grau II	Pele G1 / TGI G2	sim-pele	39	extenso	159	quiescente	474	morto	Recaída
13	85558	TMO433	14	grau II	Pele G2 /TGI G1	não	23	extenso	177	quiescente	832	vivo	----
14	85559	TMO523	12	ausência	----	----	----	extenso	175	de novo	742	vivo	----
15	86801	TMO254	16	ausência	----	----	----	extenso	168	de novo	872	vivo	----
16	85560	TMO588	20	ausência	----	----	----	extenso	200	de novo	597	vivo	----
17	85575	TMO525	13	ausência	----	----	----	extenso	202	de novo	676	morto	GVHD crônico +PTLD
18	85572	TMO219	14	grau IV	Pele G3 /Fígado G2/ TGI G2	Não	30	ausência	----	-----	35	morto	GVHD agudo+ sepse de foco pneumônico
19	85581	TMO554	11	grau II	Fígado G1/ Pele G1/TGI G1	Não	19	ausência	----	-----	209	morto	Recaída
20	85582	TMO440	13	ausência	----	----	----	ausência	----	-----	174	morto	Recaída
21	85583	TMO243	21	ausência	----	----	----	ausência	----	-----	1076	vivo	----
22	85584	TMO457	16	ausência	----	----	----	ausência	----	-----	284	morto	Recaída
23	85591	TMO240	20	grau II	Fígado G1/ Pele G1	sim-pele	40	extenso	147	quiescente	443	morto	GVHD crônico+ Neurotoxoplasmose

Número Paciente	Array	Amostra	SR	SD	IR (anos)	RC	RC ablação	Profilaxia GVHD	Doença	Status ao TMO	Risco da Doença	Fonte de célula	Incompat ABO	Incompat Sexo (F→M)
24	85596	TMO485	M	M	33	BuCy	MA	CSA+MTX	LMC	Fase Crônica	baixo	MO	ausente	Não
25	85647	TMO209	M	F	63	Bu+ATG+ Flu	RTR	CSA+MTX	SMD-AR	----	baixo	SP	menor	Sim
26	85882	TMO181	M	M	33	BuCy	MA	CSA+MTX	LMC	2ª Fase Crônica	interm	SP	ausente	Não
27	85883	TMO557	F	F	37	BuCy	MA	CSA+MTX	LMC	Fase Crônica	baixo	MO	ausente	Não
28	85648	TMO108	M	F	32	BuCy	MA	CSA+MTX	LMC	Fase Crônica	baixo	SP	maior	Sim
29	86802	TMO683	F	M	42	BuCy	MA	CSA+MTX	SMD AREB II	----	alto	SP	ausente	Não
30	85649	TMO579	F	M	48	BuCy	MA	CSA+MTX	SMD AREB II	----	alto	SP	ausente	Não
31	85650	TMO427	M	F	43	FluBu	MA	CSA+MTX	Leucemia Mielóide Crônica	Fase Crônica	baixo	SP	ausente	Sim
32	85651	TMO067	F	F	55	CyMel	RTR	CSA+MTX	LMMC	Fase Blástica	alto	SP	ausente	Não
33	85662	TMO365	M	M	30	BuCy	MA	CSA+MTX	LMC	Fase Crônica	baixo	MO	ausente	Não
34	86980	TMO620	M	M	54	TBI+Flu	RTR	CSA+MMF	LNH da Zona do Manto IIIA	Remissão	baixo	SP	ausente	Não
35	85809	TMO589	M	F	17	TBI+Cy	MA	CSA+MTX	LLA pró B	1ª Remissão completa	baixo	SP	ausente	Sim
36	85884	TMO635	M	M	37	BuCy	MA	CSA+MTX	AAG	----	não malig	MO	menor	Não
37	85889	TMO403	F	F	43	BuCy	MA	CSA+MTX	LLA Pré B	1ª Remissão completa	baixo	SP	ausente	Não
38	86804	TMO298	M	F	44	CyMel	RTR	CSA+MTX	LNH Folicular Grau II	3ª recidiva	alto	SP	ausente	Sim
39	85877	TMO607	F	F	17	BuCy	MA	CSA+MTX	AAG	----	não malig	MO	ausente	Não
40	85876	TMO093	F	M	16	BuCy	MA	CSA+MTX	SMD hipocelular	----	baixo	SP	ausente	Não
41	85875	TMO163	F	F	28	BuCy	MA	CSA+MTX	AAG	----	não malig	MO	menor	Não
42	85874	TMO526	F	M	24	BuCy	MA	CSA+MTX	LMA sem maturação (FAB-M1)	1ª Remissão completa	baixo	SP	ausente	Não
43	85873	TMO231	F	F	36	BEAM	MA	CSA+MTX	LNH Anaplásico IVB	Refratário	alto	SP	ausente	Não
44	85872	TMO685	M	M	36	BuCy	MA	CSA+MTX	LMC	Fase Acelerada	interm	MO	ausente	Não

Número Paciente	Array	Amostra	Enxertia medular(dias)	aGVHD - Graduação	aGVHD por sítio	Biópsia-aGVHD	Diagnóstico aGVHD (dias)	cGVHD extensão	Diagnóstico cGVHD (dias)	cGVHD tipo de início	Seg(dias)	Status	Causa do óbito
24	85596	TMO485	22	ausência	----	----	----	extenso	207	de novo	305	morto	GVHD crônico + Pneumonite por CMV
25	85647	TMO209	12	ausência	----	----	----	extenso	108	de novo	1125	vivo	----
26	85882	TMO181	14	grau II	Pele G1/ TGI G1	sim-pele	33	extenso	169	quiescente	1084	vivo	----
27	85883	TMO557	19	ausência	----	----	----	localizado	159	de novo	565	vivo	----
28	85648	TMO108	13	ausência	----	----	----	extenso	186	de novo	1200	vivo	----
29	86802	TMO683	14	grau II	Pele G2 /TGIG2	Não	62	ausência	----	-----	64	morto	Sepse
30	85649	TMO579	12	grau II	Pele G2/ TGI G1	sim-pele	25	extenso	306	quiescente	419	vivo	----
31	85650	TMO427	17	grau III	Pele G2/ Fígado G2/ TGIG2	sim-pele	38	extenso	113	quiescente	854	vivo	----
32	85651	TMO067	11	ausência	----	----	----	localizado	158	de novo	1151	vivo	----
33	85662	TMO365	17	ausência	----	----	----	ausência	----	-----	883	vivo	----
34	86980	TMO620	14	ausência	----	----	----	localizado	152	de novo	658	vivo	----
35	85809	TMO589	11	ausência	----	----	----	extenso	136	de novo	715	vivo	----
36	85884	TMO635	29	ausência	----	----	----	extenso	102	de novo	613	vivo	----
37	85889	TMO403	14	grau III	PeleG2 /FígadoG2 /TGI G3	Não	33	ausência	----	-----	53	morto	GVHD agudo + Sepse
38	86804	TMO298	16	grau II	Pele G3	sim-pele	35	ausência	----	-----	365	vivo	----
39	85877	TMO607	24	ausência	----	----	----	ausência	----	-----	180	morto	Perda do Enxerto + Sepse
40	85876	TMO093	10	ausência	----	----	----	localizado	750	de novo	1198	vivo	----
41	85875	TMO163	19	ausência	----	----	----	extenso	140	de novo	1055	vivo	----
42	85874	TMO526	11	ausência	----	----	----	extenso	174	de novo	744	vivo	----
43	85873	TMO231	11	grau III	Pele G3/ TGI G3	sim-pele	25	extenso	95	quiescente	863	vivo	----
44	85872	TMO685	14	ausência	----	----	----	localizado	92	de novo	486	vivo	----

Número Paciente	Array	Amostra	SR	SD	IR (anos)	RC	RC ablação	Profilaxia GVHD	Doença	Status ao TMO	Risco da Doença	Fonte de célula	Incompat ABO	Incompat Sexo (F→M)
45	85871	TMO008	F	M	41	BuCy	MA	CSA+MTX	AAG	----	não malig	MO	ausente	Não
46	85870	TMO408	F	M	17	BuCy	MA	CSA+MTX	LMA -minimamente diferenciada (FAB-M0)	Refratária	alto	SP	ausente	Não
47	85815	TMO725	M	F	39	BuCy	MA	CSA+MTX	Leucemia Mielóide Crônica	Fase Crônica	baixo	MO	ausente	Sim
48	85814	TMO711	M	M	55	TBI+Flu	RTR	CSA+MMF	Leucemia de Células Plasmáticas	Remissão parcial	alto	SP	ausente	Não
49	85813	TMO716	F	F	69	TBI+Flu	RTR	CSA+MMF	SMD AREB II	----	alto	SP	ausente	Não
50	85810	TMO741	F	M	46	TBI+Flu	RTR	CSA+MMF	Linfoma Folicular Grau II IVA	3ª Recidiva	alto	SP	ausente	Não
51	85914	TMO704	F	M	38	TBI+Cy	MA	CSA+MTX	LLA pré B	1ª Remissão completa	baixo	SP	ausente	Não
52	85915	TMO606	M	M	30	FluBu	MA	CSA+MTX	LMA -minimamente diferenciada (FAB-M0)	Refratária	alto	SP	ausente	Não
53	86243	TMO762	F	F	50	BuCy	MA	CSA+MTX	LMA minimamente diferenciada(FAB-M0)	1ª Remissão completa	baixo	SP	menor	Não
54	86320	TMO790	M	M	51	TBI+Flu	RTR	CSA+MMF	LHEN IVB	Recidiva após autólogo	alto	SP	ausente	Não
55	86800	TMO784	F	F	46	TBI+Cy	MA	CSA+MTX	LMA com maturação (FAB-M2)	Refratária	alto	SP	maior	Não
56	86798	TMO796	F	M	22	TBI+Cy	MA	CSA+MTX	LLA T	2ª Remissão completa	Interm	SP	maior	Não
57	86716	TMO794	F	M	52	TBI+Flu	RTR	CSA+MMF	MM IIIA A IgG Kappa	Recidiva após autólogo	alto	SP	menor	Não
58	86997	TMO809	F	M	36	BuCy	MA	CSA+MTX	LMA com maturação (FAB-M2)	1ª Remissão completa	baixo	SP	ausente	Não
59	86319	TMO125	M	F	32	TBI+Cy	MA	CSA+MTX	LNH Linfoblástico IVB	Remissão parcial	alto	SP	ausente	Sim
60	86318	TMO206	F	M	20	BEAM	MA	CSA+MTX	LHEN IVB	Refratário	alto	SP	ausente	Não
61	86313	TMO126	M	F	17	BEAM	MA	CSA+MTX	LHEN IVB	Refratário	alto	SP	maior	Sim
62	86312	TMO342	M	F	22	BuCy	MA	CSA+MTX	LMA -minimamente diferenciada (FAB-M0)	Refratária	alto	SP	ausente	Sim
63	86311	TMO377	F	M	18	BuCy	MA	CSA+MTX	LLA Pré B	1ª Remissão completa	baixo	SP	ausente	Não
64	86259	TMO304	F	M	53	Mel 140	RTR	CSA+MTX	MM IgG kappa IIIA	Refratário	alto	SP	ausente	Não
65	86258	TMO481	M	F	47	BuCy	MA	CSA+MTX	SMD hipocelular	----	baixo	SP	maior	Sim
66	86254	TMO662	F	F	55	TBI+Flu	RTR	CSA+MMF	Leucemia Monoblástica aguda (FAB:M5a)	1ª Remissão completa	baixo	SP	ausente	Não

Número Paciente	Array	Amostra	Enxertia medular(dias)	aGVHD - Graduação	aGVHD por sítio	Biópsia-aGVHD	Diagnóstico aGVHD (dias)	cGVHD extensão	Diagnóstico cGVHD (dias)	cGVHD tipo de início	Seg(dias)	Status	Causa do óbito
45	85871	TMO008	20	ausência	----	----	----	ausência	----	-----	1278	vivo	----
46	85870	TMO408	12	grau II	Pele G2 /TGI G1	sim-pele	50	ausência	----	-----	142	morto	Recaída
47	85815	TMO725	20	ausência	----	----	----	extenso	182	de novo	527	vivo	----
48	85814	TMO711	14	ausência	----	----	----	ausência	----	-----	231	vivo	----
49	85813	TMO716	19	grau IV	TGIG4 /FigadoG2	Não	88	ausência	----	-----	166	morto	GVHD agudo
50	85810	TMO741	13	ausência	----	----	----	extenso	221	de novo	338	morto	GVHD crônico + desnutrição protéico calórica
51	85914	TMO704	14	grau I	Pele G2	sim-pele	37	ausência	----	-----	68	morto	Aspergilose Pulmonar
52	85915	TMO606	18	ausência	----	----	----	extenso	116	de novo	162	morto	Recaída
53	86243	TMO762	12	grau II	Pele G2 /TGI G1	sim -pele	46	extenso	100	progressivo	255	vivo	----
54	86320	TMO790	15	grau II	Pele G2/ TGI G1	Não	25	extenso	167	quiescente	384	morto	Recaída
55	86800	TMO784	21	ausência	----	----	----	extenso	125	de novo	415	vivo	----
56	86798	TMO796	15	grau III	TGI G3 /PeleG3	Não	80	ausência	----	-----	98	morto	GVHD agudo +Sepse
57	86716	TMO794	14	ausência	----	----	----	extenso	210	de novo	246	morto	BOOP
58	86997	TMO809	14	ausência	----	----	----	ausência	----	-----	164	morto	Recaída
59	86319	TMO125	10	ausência	----	----	----	extenso	247	de novo	1184	vivo	----
60	86318	TMO206	10	grau II	Fígado G1/ TGI G1	----	29	extenso	261	quiescente	1095	vivo	----
61	86313	TMO126	13	ausência	----	----	----	extenso	151	de novo	1184	vivo	----
62	86312	TMO342	13	grau III	TGI G3 /Figado G3 /PeleG2	Não	74	ausência	----	-----	81	morto	GVHD agudo
63	86311	TMO377	14	ausência	----	----	----	ausência	----	-----	150	morto	Recaída
64	86259	TMO304	13	ausência	----	----	----	extenso	151	de novo	238	morto	GVHD crônico + Sepse
65	86258	TMO481	18	grau IV	Pele G4/ TGI 4/ Figado2	Não	36	ausência	----	-----	44	morto	GVHD agudo
66	86254	TMO662	12	grau II	PeleG2/ TGI2	sim-pele	22	extenso	155	quiescente	396	morto	Recaída

Número Paciente	Array	Amostra	SR	SD	IR (anos)	RC	RC ablação	Profilaxia GVHD	Doença	Status ao TMO	Risco da Doença	Fonte de célula	Incompat ABO	Incompat Sexo (F→M)
67	86253	TMO483	F	F	48	BuCy	MA	CSA+MTX	LMA com displasia multilinhagem	1ª Remissão completa	alto	SP	ausente	Não
68	86252	TMO363	F	M	47	BuCy	MA	CSA+MTX	SMD hipocelular	-----	baixo	SP	ausente	Não
69	87000	TMO007	F	F	38	TBI+Cy	MA	CSA+MTX	LBA monocítica/B	Refratária	alto	SP	ausente	Não
70	87001	TMO482	F	M	50	FluBu	MA	CSA+MTX	SMD AREB II	-----	alto	SP	ausente	Não
71	86791	TMO654	F	F	41	BuCy	MA	CSA+MTX	LMA com maturação (FAB-M2)	1ª Remissão completa	baixo	SP	ausente	Não
72	86790	TMO045	M	F	31	BuCy	MA	CSA+MTX	LMC	Fase Crônica	baixo	MO	ausente	Sim
73	86789	TMO112	F	F	28	BuCy	MA	CSA+MTX	LMC	Fase Acelerada	interm	SP	ausente	Não
74	86787	TMO295	M	M	46	BuCy	MA	CSA+MTX	LMC	Fase Crônica	baixo	MO	maior	Não
75	86773	TMO653	M	M	60	TBI+Flu	RTR	CSA+MMF	LMA com maturação (FAB-M2)	1ª Remissão completa	baixo	SP	maior	Não
76	86772	TMO091	F	F	48	BuCy	MA	CSA+MTX	LMA com maturação (FAB-M2)	2ª Remissão completa	Interm	SP	ausente	Não
77	86709	TMO717	M	F	21	BuCy	MA	CSA+MTX	AAG	-----	não malig	MO	ausente	Sim
78	86699	TMO386	F	F	46	BEAM	MA	CSA+MTX	LHEN IVB	Refratário	alto	MO	ausente	Não
79	86330	TMO336	M	F	19	BuCy	MA	CSA+MTX	LNH Linfoblastico T IVB	Refratário	alto	SP	ausente	Sim
80	86981	TMO092	M	M	48	Bu+ATG+ Flu	RTR	CSA+MTX	MM IgA lambda IIIB	Refratário	alto	SP	menor	Não
81	86983	TMO760	F	F	17	BuCy	MA	CSA+MTX	AAG	-----	não malig	MO	ausente	Não
82	86984	TMO783	F	M	30	BuCy	MA	CSA+MTX	Leucemia Monoblástica aguda (FAB:M5a)	1ª Remissão completa	baixo	SP	ausente	Não
83	86982	TMO804	F	M	21	BEAM	MA	CSA+MTX	LHEN IVB	Refratário	alto	SP	menor	Não
84	86714	TMO808	F	F	57	TBI+Flu	RTR	CSA+MMF	LNH Folicular Grau II	3ª recidiva	alto	SP	ausente	Não
85	86323	TMO822	M	F	19	BuCy	MA	CSA+MTX	LMC	Fase Crônica	baixo	MO	menor	Sim
86	86324	TMO821	M	M	52	TBI+Flu	RTR	CSA+MMF	LMC	Fase Crônica	baixo	SP	ausente	Não
87	86325	TMO815	M	M	41	BuCy	MA	CSA+MTX	LMA com maturação (FAB-M2)	1ª Remissão completa	baixo	SP	ausente	Não
88	86998	TMO836	M	M	45	BuCy	MA	CSA+MTX	LMC	Fase Crônica	baixo	MO	ausente	Não
89	86999	TMO837	F	M	51	BuCy	MA	CSA+MTX	Leucemia Mielomonocítica Aguda (FAB-M4)	1ª Remissão completa	baixo	SP	ausente	Não

Número Paciente	Array	Amostra	Enxertia medular(dias)	aGVHD - Graduação	aGVHD por sítio	Biópsia-aGVHD	Diagnóstico aGVHD (dias)	cGVHD extensão	Diagnóstico cGVHD (dias)	cGVHD tipo de início	Seg(dias)	Status	Causa do óbito
67	86253	TMO483	11	grau III	Fígado G2/ TGI2	----	35	ausência	----	-----	101	morto	GVHD agudo + Artrite Séptica
68	86252	TMO363	13	ausência	----	----	----	extenso	261	de novo	897	vivo	----
69	87000	TMO007	15	grau II	Pele G1/ TGI G1	Não	32	ausência	----	-----	160	morto	Recaída
70	87001	TMO482	19	ausência	----	----	----	extenso	200	de novo	211	morto	GVHD crônico + Sepse
71	86791	TMO654	12	grau III	Fígado G2	sim-hepática	20	extenso	108	quiescente	593	vivo	----
72	86790	TMO045	15	grau III	Fígado G2/ TGI G2	não	24	extenso	192	progressivo	210	morto	GVHD crônico + Neurotoxoplasmose
73	86789	TMO112	14	ausência	----	----	----	ausência	----	-----	199	morto	Sepse
74	86787	TMO295	19	ausência	----	----	----	extenso	106	de novo	1012	vivo	----
75	86773	TMO653	14	ausência	----	----	----	ausência	----	-----	610	vivo	----
76	86772	TMO091	13	ausência	----	----	----	ausência	----	-----	1204	vivo	----
77	86709	TMO717	19	grau I	Pele G2	sim-pele	38	extenso	113	quiescente	417	vivo	----
78	86699	TMO386	15	grau III	TGI G3/ Fígado G3	sim-intestino	40	ausência	----	-----	41	morto	GVHD agudo + Sepse
79	86330	TMO336	13	grau III	TGI G2/ Pele G2	sim-pele	40	ausência	----	-----	59	morto	GVHD agudo + CMV
80	86981	TMO092	13	ausência	----	----	----	extenso	121	de novo	1017	vivo	----
81	86983	TMO760	18	ausência	----	----	----	ausência	----	-----	511	vivo	----
82	86984	TMO783	13	ausência	----	----	----	localizado	155	de novo	363	vivo	----
83	86982	TMO804	13	grau II	TGI G1/Pele G2/ Fígado G1	Não	21	ausência	----	-----	343	vivo	----
84	86714	TMO808	16	grau II	Pele G2 /Fígado G1	Não	30	ausência	----	-----	127	morto	Pneumonia Intersticial
85	86323	TMO822	17	ausência	----	----	----	ausência	----	-----	326	vivo	----
86	86324	TMO821	19	ausência	----	----	----	extenso	102	de novo	308	vivo	----
87	86325	TMO815	11	ausência	----	----	----	extenso	171	de novo	318	vivo	----
88	86998	TMO836	17	grau II	Pele G2/ TGI G1	sim-intestinal	33	ausência	----	-----	150	morto	Pneumonia Estafilocócica
89	86999	TMO837	14	grau I	Pele G1	sim-pele	25	ausência	----	-----	266	vivo	----

SR: sexo do receptor ; SD : sexo do doador; IR: idade do receptor; RC: regime de condicionamento; aGVHD: doença do enxerto versus hospedeiro aguda; cGVHD: doença do enxerto versus hospedeiro crônica; MA: mieloablativo; RTR : regimede toxicidade reduzida; Bu: bussulfano; Cy: ciclofosfamida; TBI: irradiação corporal total; Flu: fludarabina; Mel : melfalan : ATG: linfoglobulina; BEAM : carmustina, etoposide, citarabina e melfalan; CSA: ciclosporina; MTX: metotrexate; MMF : micofenolato mofetil; SP: células periféricas; MO: medula óssea; AAG –Anemia Aplástica Grave; HPN – Hemoglobinúria Paroxística Noturna; LBA- Leucemia Bifenotípica; LH –Linfoma de Hodgkin ; LLA –Leucemia Linfóide aguda; LMA –Leucemia Mielóide Aguda; LMC-Leucemia Mielóide Crônica; LNH – Linfoma Não Hodgkin; MM-Mieloma Múltiplo; SMD – Síndrome Mielodisplásica ; F→M: transplante de doador sexo feminino para receptor masculino; GVHD agudo por sítio: ``G'' significa a graduação por sítio acometido pela doença do enxerto aguda ; CMV : citomegalov

Anexo 3. Composição do Classificador de 233 genes

Clone	Símbolo do Gene	Descrição
<u>H200000646</u>	<u>DEFA4</u>	defensin, alpha 4, corticostatin (DEFA4), mRNA.
<u>H200005875</u>	<u>RNASE3</u>	ribonuclease, RNase A family, 3 (eosinophil cationic protein) (RNASE3), mRNA.
<u>H200011085</u>	<u>CTSG</u>	cathepsin G (CTSG), mRNA.
<u>H200011001</u>	<u>ELA2</u>	elastase 2, neutrophil (ELA2), mRNA.
<u>H200006243</u>	<u>HPR</u>	haptoglobin-related protein (HPR), mRNA.
<u>H200000469</u>	<u>MPO</u>	myeloperoxidase (MPO), nuclear gene encoding mitochondrial protein, mRNA.
<u>H200019509</u>	<u>BEX1</u>	brain expressed, X-linked 1 (BEX1), mRNA.
<u>H200009720</u>	<u>GDF15</u>	growth differentiation factor 15 (GDF15), mRNA.
<u>H200009821</u>	<u>MYH10</u>	myosin, heavy chain 10, non-muscle (MYH10), mRNA.
<u>H200011744</u>	<u>ALOX15B</u>	arachidonate 15-lipoxygenase, type B (ALOX15B), transcript variant b, mRNA.
<u>H200008642</u>	<u>MAOA</u>	monoamine oxidase A (MAOA), nuclear gene encoding mitochondrial protein, mRNA.
<u>H200004906</u>	<u>CDCA8</u>	cell division cycle associated 8 (CDCA8), mRNA.
<u>H200015984</u>	<u>HBO1</u>	hemoglobin, theta 1 (HBO1), mRNA.
<u>H200011257</u>	<u>RIPK2</u>	receptor-interacting serine-threonine kinase 2 (RIPK2), mRNA.
<u>H200000275</u>	<u>PSCD1</u>	pleckstrin homology, Sec7 and coiled-coil domains 1 (cytohesin 1) (PSCD1), transcript variant 2, mRNA.
<u>H200011281</u>	<u>Igk-V28</u>	Immunoglobulin kappa chain variable 28 (V28)
<u>H200000316</u>	<u>ADH4</u>	alcohol dehydrogenase 4 (class II), pi polypeptide (ADH4), mRNA.
<u>H200014737</u>	<u>PTGS2</u>	prostaglandin-endoperoxide synthase 2 (prostaglandin G/H synthase and cyclooxygenase) (PTGS2)
<u>H200007941</u>	<u>TMOD1</u>	tropomodulin 1 (TMOD1), mRNA.
<u>H200014027</u>	<u>CA2</u>	carbonic anhydrase II (CA2), mRNA.
<u>H200005476</u>	<u>F3</u>	coagulation factor III (thromboplastin, tissue factor) (F3), mRNA.
<u>H200017397</u>	<u>PURG</u>	purine-rich element binding protein G (PURG), transcript variant A, mRNA.
<u>H200009102</u>	<u>FLJ21272</u>	Hypothetical protein FLJ21272
<u>H200017533</u>	<u>THG1L</u>	TRNA-histidine guanylyltransferase 1-like (<i>S. cerevisiae</i>)
<u>H200006961</u>	<u>HMBS</u>	hydroxymethylbilane synthase (HMBS), transcript variant 2, mRNA.
<u>H200002660</u>	<u>C11orf1</u>	Chromosome 11 open reading frame 1
<u>H200015643</u>	<u>PBEF1</u>	pre-B-cell colony enhancing factor 1 (PBEF1), mRNA.
<u>H200014292</u>	<u>STK38L</u>	serine/threonine kinase 38 like (STK38L), mRNA.

H200019041	ALAS2	aminolevulinate, delta-, synthase 2 (sideroblastic/hypochromic anemia) (ALAS2), nuclear gene encoding mitochondrial protein, transcript variant 4, mRNA.
H200004352	HBD	hemoglobin, delta (HBD), mRNA.
H200016970	LOC647979	PREDICTED: hypothetical LOC647979, transcript variant 2 (LOC647979), mRNA.
H200014600	GOSR1	golgi SNAP receptor complex member 1 (GOSR1), transcript variant 2, mRNA.
H200014822		MRNA full length insert cDNA clone EUROIMAGE 1895238
H200014448	TMEM176B	transmembrane protein 176B (TMEM176B), mRNA.
H200010342	BPI	bactericidal/permeability-increasing protein (BPI), mRNA.
H200020633		CDNA FLJ25041 fis, clone CBL03194
H200012248	ADAMTS2	ADAM metalloproteinase with thrombospondin type 1 motif, 2 (ADAMTS2), transcript variant 1, mRNA.
H200018390	TRSPAP1	TRNA selenocysteine associated protein 1
H200017053	ERAF	erythroid associated factor (ERAF), mRNA.
H200013038	FBXL4	F-box and leucine-rich repeat protein 4
H200017694	PCDHB6	protocadherin beta 6 (PCDHB6), mRNA.
H200014442	TRIM65	tripartite motif-containing 65 (TRIM65), mRNA.
H200014375	ZNF524	zinc finger protein 524 (ZNF524), mRNA.
H200009119	ZFP2	zinc finger protein 2 homolog (mouse) (ZFP2), mRNA.
H200015022	ITIH5	inter-alpha (globulin) inhibitor H5 (ITIH5), transcript variant 2, mRNA.
H200007969		CDNA: FLJ22110 fis, clone HEP18112
H200009355	METTL7A	methyltransferase like 7A (METTL7A), mRNA.
H200019974	RAD51	RAD51 homolog (RecA homolog, E. coli) (S. cerevisiae) (RAD51), transcript variant 2, mRNA.
H200016982	AGRN	agrin (AGRN), mRNA.
H200005341	HHAT	hedgehog acyltransferase (HHAT), mRNA.
H200018080	ELOVL3	elongation of very long chain fatty acids (FEN1/Elo2, SUR4/Elo3, yeast)-like 3 (ELOVL3), mRNA.
H200007119	KLF5	Kruppel-like factor 5 (intestinal) (KLF5), mRNA.
H200019894	CACNA2D3	calcium channel, voltage-dependent, alpha 2/delta 3 subunit (CACNA2D3), mRNA.
H200011529	SRFBP1	serum response factor binding protein 1 (SRFBP1), mRNA.
H200017296	DKK3	Dickkopf homolog 3 (Xenopus laevis)
H200016988	ANAPC4	Anaphase promoting complex subunit 4
H200012291	C21orf55	chromosome 21 open reading frame 55 (C21orf55), transcript variant 2, mRNA.
H200015094	ATXN7L1	Ataxin 7-like 1
H200015250	FLJ40244	Hypothetical protein FLJ40244
H200002607	SPIRE1	spire homolog 1 (Drosophila) (SPIRE1), mRNA.

<u>H200017662</u>	<u>PRO2964</u>	Hypothetical protein PRO2964
<u>H200014723</u>		CDNA FLJ12049 fis, clone HEMBB1001996
<u>H200015350</u>	<u>DDX19A</u>	DEAD (Asp-Glu-Ala-As) box polypeptide 19A (DDX19A), mRNA.
<u>H200003497</u>	<u>PRLR</u>	Prolactin receptor
<u>H200002119</u>	<u>ADCY6</u>	adenylate cyclase 6 (ADCY6), transcript variant 2, mRNA.
<u>H200014086</u>	<u>CHST10</u>	carbohydrate sulfotransferase 10 (CHST10), mRNA.
<u>H200011322</u>	<u>FLJ10292</u>	mago-nashi homolog 2 (FLJ10292), mRNA.
<u>H200017362</u>	<u>IL1F6</u>	interleukin 1 family, member 6 (epsilon) (IL1F6), mRNA.
<u>H200018605</u>		CDNA FLJ25271 fis, clone STM05584
<u>H200019262</u>		CDNA FLJ11723 fis, clone HEMBA1005314
<u>H200006346</u>	<u>SNCA</u>	synuclein, alpha (non A4 component of amyloid precursor) (SNCA), transcript variant NACP112, mRNA.
<u>H200017000</u>	<u>EPB49</u>	erythrocyte membrane protein band 4.9 (dematin) (EPB49), mRNA.
<u>H200019373</u>	<u>MGC13219</u>	Hypothetical protein MGC13219
<u>H200014134</u>	<u>FAM53B</u>	family with sequence similarity 53, member B (FAM53B), mRNA.
<u>H200010278</u>	<u>BACH2</u>	BTB and CNC homology 1, basic leucine zipper transcription factor 2 (BACH2), mRNA.
<u>H200006587</u>	<u>RGS2</u>	regulator of G-protein signalling 2, 24kDa (RGS2), mRNA.
<u>H200012604</u>	<u>IL1B</u>	interleukin 1, beta (IL1B), mRNA.
<u>H200017325</u>	<u>IFI27</u>	Interferon, alpha-inducible protein 27
<u>H200012695</u>		Transcribed locus
<u>H200012031</u>	<u>OTC</u>	Ornithine carbamoyltransferase
<u>H200000451</u>	<u>IL1A</u>	interleukin 1, alpha (IL1A), mRNA.
<u>H200013727</u>	<u>NDUFS7</u>	NADH dehydrogenase (ubiquinone) Fe-S protein 7, 20kDa (NADH-coenzyme Q reductase)
<u>H200014514</u>	<u>ADARB2</u>	adenosine deaminase, RNA-specific, B2 (RED2 homolog rat) (ADARB2), mRNA.
<u>H200005869</u>	<u>ABCB4</u>	ATP-binding cassette, sub-family B (MDR/TAP), member 4 (ABCB4), transcript variant B, mRNA.
<u>H200001408</u>	<u>SFRP1</u>	secreted frizzled-related protein 1 (SFRP1), mRNA.
<u>H200012113</u>		Transcribed locus
<u>H200005171</u>	<u>CAPNS2</u>	calpain, small subunit 2 (CAPNS2), mRNA.
<u>H200012645</u>	<u>UPB1</u>	ureidopropionase, beta (UPB1), mRNA.
<u>H200011582</u>	<u>THBS2</u>	Thrombospondin 2
<u>H200010246</u>	<u>ALDH3B2</u>	Aldehyde dehydrogenase 3 family, member B2
<u>H200000028</u>	<u>CYP4F2</u>	cytochrome P450, family 4, subfamily F, polypeptide 2 (CYP4F2), mRNA.
<u>H200010979</u>	<u>SNIP1</u>	Smad nuclear interacting protein 1 (SNIP1), mRNA.
<u>H200008840</u>	<u>C12orf30</u>	chromosome 12 open reading frame 30 (C12orf30), mRNA.

<u>H200000684</u>	<u>POU6F1</u>	POU domain, class 6, transcription factor 1 (POU6F1), mRNA.
<u>H200014118</u>	<u>CREM</u>	cAMP responsive element modulator (CREM), transcript variant 21, mRNA.
<u>H200010352</u>	<u>ETV1</u>	Ets variant gene 1
<u>H200013469</u>	<u>E2F7</u>	E2F transcription factor 7 (E2F7), mRNA.
<u>H200018501</u>	<u>RAB4B</u>	RAB4B, member RAS oncogene family
<u>H200014865</u>	<u>CD207</u>	CD207 molecule, langerin (CD207), mRNA.
<u>H200006520</u>	<u>GAS6</u>	growth arrest-specific 6 (GAS6), mRNA.
<u>H200008742</u>	<u>HIBADH</u>	3-hydroxyisobutyrate dehydrogenase
<u>H200013312</u>	<u>CAV2</u>	caveolin 2 (CAV2), transcript variant 2, mRNA.
<u>H200021244</u>	<u>LOC645158</u>	PREDICTED: hypothetical protein LOC645158 (LOC645158), mRNA.
<u>H200006773</u>	<u>NR4A3</u>	nuclear receptor subfamily 4, group A, member 3 (NR4A3), transcript variant 3, mRNA.
<u>H200006911</u>	<u>GPNMB</u>	glycoprotein (transmembrane) nmb (GPNMB), transcript variant 2, mRNA.
<u>H200017313</u>	<u>FGFR2</u>	Fibroblast growth factor receptor 2 (bacteria-expressed kinase, keratinocyte growth factor receptor, craniofacial dysostosis 1, Crouzon syndrome, Pfeiffer syndrome, Jackson-Weiss syndrome)
<u>H200000507</u>	<u>SPTA1</u>	Spectrin, alpha, erythrocytic 1 (elliptocytosis 2)
<u>H200010953</u>	<u>ZNF659</u>	zinc finger protein 659 (ZNF659), mRNA.
<u>H200015790</u>	<u>PLEKHA5</u>	Pleckstrin homology domain containing, family A member 5
<u>H200015220</u>	<u>G3BP1</u>	GTPase activating protein (SH3 domain) binding protein 1 (G3BP1), transcript variant 2, mRNA.
<u>H200014721</u>	<u>SP8</u>	Sp8 transcription factor (SP8), transcript variant 1, mRNA.
<u>H200005086</u>	<u>CRISP3</u>	cysteine-rich secretory protein 3 (CRISP3), mRNA.
<u>H200016920</u>	<u>CASR</u>	calcium-sensing receptor (hypocalciuric hypercalcemia 1, severe neonatal hyperparathyroidism) (CASR), mRNA.
<u>H200021246</u>	<u>DTWD1</u>	DTW domain containing 1
<u>H200004325</u>	<u>ALPK3</u>	alpha-kinase 3 (ALPK3), mRNA.
<u>H200003661</u>	<u>TMEM161A</u>	transmembrane protein 161A (TMEM161A), mRNA.
<u>H200009424</u>	<u>FLJ20920</u>	hypothetical protein FLJ20920 (FLJ20920), mRNA.
<u>H200012911</u>	<u>ZNF180</u>	zinc finger protein 180 (ZNF180), mRNA.
<u>H200008702</u>		Transcribed locus
<u>H200010515</u>	<u>ARID4A</u>	AT rich interactive domain 4A (RBP1-like)
<u>H200021216</u>		CDNA: FLJ21543 fis, clone COL06171
<u>H200012785</u>	<u>ODF3</u>	Outer dense fiber of sperm tails 3
<u>H200001060</u>	<u>EXDL2</u>	exonuclease 3'-5' domain-like 2 (EXDL2), mRNA.
<u>H200004578</u>	<u>MSC</u>	Musculin (activated B-cell factor-1)
<u>H200019074</u>	<u>C1orf77</u>	chromosome 1 open reading frame 77 (C1orf77), mRNA.

<u>H200010020</u>	<u>CABP1</u>	calcium binding protein 1 (CABP1), transcript variant 1, mRNA.
<u>H200007927</u>		CDNA FLJ11519 fis, clone HEMBA1002348
<u>H200002050</u>	<u>EPPB9</u>	B9 protein (EPPB9), mRNA.
<u>H200015633</u>	<u>SLC16A9</u>	solute carrier family 16, member 9 (monocarboxylic acid transporter 9) (SLC16A9), mRNA.
<u>H200015554</u>	<u>TBCCD1</u>	TBCC domain containing 1 (TBCCD1), mRNA.
<u>H200013901</u>	<u>NEK2</u>	NIMA (never in mitosis gene a)-related kinase 2 (NEK2), mRNA.
<u>H200016084</u>	<u>KCNA10</u>	potassium voltage-gated channel, shaker-related subfamily, member 10 (KCNA10), mRNA.
<u>H200013302</u>	<u>PDE5A</u>	phosphodiesterase 5A, cGMP-specific (PDE5A), transcript variant 3, mRNA.
<u>H200003615</u>	<u>LINGO2</u>	leucine rich repeat and Ig domain containing 2 (LINGO2), mRNA.
<u>H200020557</u>		PRO1787
<u>H200013509</u>	<u>ASPM</u>	asp (abnormal spindle) homolog, microcephaly associated (Drosophila) (ASPM), mRNA.
<u>H200004857</u>	<u>ARHGEF11</u>	Rho guanine nucleotide exchange factor (GEF) 11 (ARHGEF11), transcript variant 1, mRNA.
<u>H200012740</u>	<u>ZNF498</u>	zinc finger protein 498 (ZNF498), mRNA.
<u>H200005118</u>	<u>JAM2</u>	junctional adhesion molecule 2 (JAM2), mRNA.
<u>H200007329</u>	<u>FPRL2</u>	Formyl peptide receptor-like 2
<u>H200000128</u>	<u>RAB5C</u>	RAB5C, member RAS oncogene family (RAB5C), transcript variant 2, mRNA.
<u>H200020117</u>	<u>MTMR1</u>	myotubularin related protein 1 (MTMR1), mRNA.
<u>H200009918</u>	<u>IGHG1</u>	Immunoglobulin heavy constant gamma 1 (G1m marker)
<u>H200016801</u>	<u>KLF15</u>	Kruppel-like factor 15 (KLF15), mRNA.
<u>H200013892</u>	<u>MLL2</u>	myeloid/lymphoid or mixed-lineage leukemia 2 (MLL2), mRNA.
<u>H200013248</u>	<u>FLJ23577</u>	KPL2 protein (FLJ23577), transcript variant 1, mRNA.
<u>H200010620</u>	<u>IL6</u>	interleukin 6 (interferon, beta 2) (IL6), mRNA.
<u>H200015466</u>	<u>C17orf73</u>	PREDICTED: chromosome 17 open reading frame 73 (C17orf73), misc RNA.
<u>H200002181</u>	<u>CUTL2</u>	cut-like 2 (Drosophila) (CUTL2), mRNA.
<u>H200010025</u>	<u>ZMYND11</u>	zinc finger, MYND domain containing 11 (ZMYND11), transcript variant 1, mRNA.
<u>H200012539</u>	<u>EPHB2</u>	EPH receptor B2 (EPHB2), transcript variant 1, mRNA.
<u>H200012511</u>	<u>MARVELD2</u>	MARVEL domain containing 2 (MARVELD2), transcript variant 2, mRNA.
<u>H200011103</u>	<u>TCF3</u>	transcription factor 3 (E2A immunoglobulin enhancer binding factors E12/E47) (TCF3), mRNA.
<u>H200007835</u>	<u>RHAG</u>	Rh-associated glycoprotein (RHAG), mRNA.
<u>H200009578</u>	<u>MCFD2</u>	multiple coagulation factor deficiency 2 (MCFD2), mRNA.
<u>H200020015</u>		Clone HQ0452 PRO0452
<u>H200014776</u>	<u>CAMK2N1</u>	calcium/calmodulin-dependent protein kinase II inhibitor 1 (CAMK2N1), mRNA.
<u>H200000477</u>	<u>MPP1</u>	membrane protein, palmitoylated 1, 55kDa (MPP1), mRNA.

<u>H200014380</u>	<u>POLR1E</u>	Polymerase (RNA) I polypeptide E, 53kDa
<u>H200018203</u>	<u>WNT5B</u>	wingless-type MMTV integration site family, member 5B (WNT5B), transcript variant 2, mRNA.
<u>H200012914</u>	<u>ADCK4</u>	aarF domain containing kinase 4 (ADCK4), mRNA.
<u>H200014007</u>	<u>ACSL1</u>	acyl-CoA synthetase long-chain family member 1 (ACSL1), mRNA.
<u>H200007924</u>	<u>AAK1</u>	AP2 associated kinase 1 (AAK1), mRNA.
<u>H200013423</u>	<u>LRRC45</u>	leucine rich repeat containing 45 (LRRC45), mRNA.
<u>H200011323</u>	<u>EIF5A2</u>	eukaryotic translation initiation factor 5A2 (EIF5A2), mRNA.
<u>H200020183</u>	<u>BPTF</u>	Bromodomain PHD finger transcription factor
<u>H200006991</u>	<u>TFF3</u>	trefoil factor 3 (intestinal) (TFF3), mRNA.
<u>H200007061</u>	<u>DHFR</u>	dihydrofolate reductase (DHFR), mRNA.
<u>H200010677</u>	<u>C1QC</u>	complement component 1, q subcomponent, C chain (C1QC), mRNA.
<u>H200010228</u>	<u>HRK</u>	harakiri, BCL2 interacting protein (contains only BH3 domain) (HRK), mRNA.
<u>H200006330</u>	<u>PSPHL</u>	Phosphoserine phosphatase-like
<u>H200015331</u>	<u>LRIT1</u>	leucine-rich repeat, immunoglobulin-like and transmembrane domains 1 (LRIT1), mRNA.
<u>H200018765</u>	<u>SLC10A7</u>	solute carrier family 10 (sodium/bile acid cotransporter family), member 7 (SLC10A7), transcript variant 3, mRNA.
<u>H200017051</u>	<u>TRIM10</u>	tripartite motif-containing 10 (TRIM10), transcript variant 2, mRNA.
<u>H200008241</u>	<u>PAWR</u>	PRKC, apoptosis, WT1, regulator (PAWR), mRNA.
<u>H200002047</u>	<u>NR0B2</u>	nuclear receptor subfamily 0, group B, member 2 (NR0B2), mRNA.
<u>H200011855</u>	<u>C3orf32</u>	chromosome 3 open reading frame 32 (C3orf32), mRNA.
<u>H200013035</u>	<u>PCYT1B</u>	phosphate cytidylyltransferase 1, choline, beta (PCYT1B), mRNA.
<u>H200018507</u>	<u>NF1</u>	Neurofibromin 1 (neurofibromatosis, von Recklinghausen disease, Watson disease)
<u>H200020374</u>	<u>MPP4</u>	Membrane protein, palmitoylated 4 (MAGUK p55 subfamily member 4)
<u>H200017676</u>		CDNA FLJ11417 fis, clone HEMBA1000960
<u>H200013024</u>	<u>MERTK</u>	C-mer proto-oncogene tyrosine kinase
<u>H200000062</u>	<u>ITGA9</u>	integrin, alpha 9 (ITGA9), mRNA.
<u>H200014896</u>	<u>LOC150005</u>	Hypothetical protein LOC150005
<u>H200004254</u>	<u>ABCC13</u>	ATP-binding cassette, sub-family C (CFTR/MRP), member 13 (ABCC13) on chromosome 21.
<u>H200020917</u>	<u>HDAC8</u>	Histone deacetylase 8
<u>H200003915</u>	<u>TNFAIP6</u>	tumor necrosis factor, alpha-induced protein 6 (TNFAIP6), mRNA.
<u>H200011018</u>	<u>AR</u>	androgen receptor (dihydrotestosterone receptor; testicular feminization; spinal and bulbar muscular atrophy; Kennedy disease) (AR), transcript variant 2, mRNA.
<u>H200018921</u>	<u>OSBPL6</u>	oxysterol binding protein-like 6 (OSBPL6), transcript variant 2, mRNA.
<u>H200015834</u>	<u>C8orf51</u>	Chromosome 8 open reading frame 51

<u>H200016959</u>	<u>CXorf48</u>	chromosome X open reading frame 48 (CXorf48), transcript variant 2, mRNA.
<u>H200011192</u>	<u>WDR40A</u>	WD repeat domain 40A (WDR40A), mRNA.
<u>H200008798</u>	<u>C6orf86</u>	Chromosome 6 open reading frame 86
<u>H200008323</u>	<u>C10orf104</u>	chromosome 10 open reading frame 104 (C10orf104), mRNA.
<u>H200002738</u>	<u>CHCHD8</u>	coiled-coil-helix-coiled-coil-helix domain containing 8 (CHCHD8), mRNA.
<u>H200020168</u>	<u>MSH5</u>	MutS homolog 5 (E. coli)
<u>H200017177</u>		MRNA; cDNA DKFZp434O0212 (from clone DKFZp434O0212)
<u>H200014964</u>	<u>PCDHB7</u>	protocadherin beta 7 (PCDHB7), mRNA.
<u>H200004154</u>	<u>HMG20B</u>	high-mobility group 20B (HMG20B), mRNA.
<u>H200002671</u>	<u>HCG9</u>	HLA complex group 9
<u>H200000678</u>	<u>KRT17</u>	keratin 17 (KRT17), mRNA.
<u>H200006588</u>	<u>CUL3</u>	cullin 3 (CUL3), mRNA.
<u>H200014029</u>	<u>EIF1AY</u>	eukaryotic translation initiation factor 1A, Y-linked (EIF1AY), mRNA.
<u>H200017919</u>		CDNA FLJ11849 fis, clone HEMBA1006709
<u>H200009273</u>	<u>KIR3DX1</u>	killer cell immunoglobulin-like receptor, three domains, X1 (KIR3DX1), mRNA.
<u>H200017938</u>		CDNA FLJ13453 fis, clone PLACE1003205
<u>H200002453</u>	<u>SRPX</u>	sushi-repeat-containing protein, X-linked (SRPX), mRNA.
<u>H200015164</u>	<u>CNTNAP3</u>	contactin associated protein-like 3 (CNTNAP3), mRNA.
<u>H200005952</u>	<u>PRIM2</u>	primase, DNA, polypeptide 2 (58kDa) (PRIM2), mRNA.
<u>H200010808</u>	<u>C4orf17</u>	chromosome 4 open reading frame 17 (C4orf17), mRNA.
<u>H200014857</u>	<u>EPN3</u>	epsin 3 (EPN3), mRNA.
<u>H200017002</u>	<u>HIGD1B</u>	HIG1 domain family, member 1B (HIGD1B), mRNA.
<u>H200002068</u>	<u>MMAB</u>	methylmalonic aciduria (cobalamin deficiency) cb1B type (MMAB), nuclear gene encoding mitochondrial protein, mRNA.
<u>H200005836</u>	<u>GNGT1</u>	guanine nucleotide binding protein (G protein), gamma transducing activity polypeptide 1 (GNGT1), mRNA.
<u>H200007530</u>	<u>MMP27</u>	matrix metalloproteinase 27 (MMP27), mRNA.
<u>H200004548</u>	<u>KRT32</u>	keratin 32 (KRT32), mRNA.
<u>H200003964</u>	<u>ZNF518</u>	zinc finger protein 518 (ZNF518), mRNA.
<u>H200006144</u>	<u>SERPINC1</u>	serpin peptidase inhibitor, clade C (antithrombin), member 1 (SERPINC1), mRNA.
<u>H200006016</u>	<u>SH3GL2</u>	SH3-domain GRB2-like 2 (SH3GL2), mRNA.
<u>H200018651</u>		Randomized negative control
<u>H200013760</u>	<u>KIAA1524</u>	KIAA1524 (KIAA1524), mRNA.
<u>H200018679</u>	<u>SPON2</u>	Spondin 2, extracellular matrix protein

<u>H2NC000004</u>		CDNA: FLJ23170 fis, clone LNG09984
<u>H200015358</u>		CDNA FLJ34007 fis, clone FCBBF1000349
<u>H200018289</u>		Transcribed locus
<u>H200011272</u>	<u>MUC7</u>	mucin 7, secreted (MUC7), mRNA.
<u>H200002866</u>	<u>NDST3</u>	N-deacetylase/N-sulfotransferase (heparan glucosaminyl) 3 (NDST3), mRNA.
<u>H200007536</u>	<u>KALRN</u>	kalirin, RhoGEF kinase (KALRN), transcript variant 3, mRNA.
<u>H200010912</u>	<u>DPY19L1P1</u>	Dpy-19-like 1 pseudogene 1 (C. elegans)
<u>H200020668</u>		Homo sapiens cDNA FLJ32951 fis, clone TESTI2008033
<u>H200014187</u>	<u>TCF7L2</u>	Transcription factor 7-like 2 (T-cell specific, HMG-box)
<u>H200013920</u>	<u>NT5E</u>	5'-nucleotidase, ecto (CD73) (NT5E), mRNA.
<u>H200015572</u>	<u>C7orf27</u>	Chromosome 7 open reading frame 27

ANEXO 4: Desempenho do classificador durante validação cruzada no grupo Treino (67 amostras)

Número Paciente	Array	Classe	Número de genes no classificador	Compound Covariate Predictor	Diagonal Linear Discriminant Analysis	1- Nearest Neighbor	1- Nearest Neighbor	Nearest Centroid	Support Vector Machines	Bayesian Compound Covariate Predictor	Classificação da amostra
21	85871	0	149	NÃO	NÃO	NÃO	NÃO	NÃO	NÃO	IC	erro
57	85572	1	137	NÃO	NÃO	NÃO	NÃO	NÃO	SIM	NÃO	erro
76	86254	1	89	NÃO	NÃO	NÃO	NÃO	NÃO	NÃO	NÃO	erro
17	85876	0	138	NÃO	NÃO	NÃO	NÃO	NÃO	SIM	NÃO	erro
61	85649	1	171	NÃO	NÃO	SIM	NÃO	NÃO	NÃO	NÃO	erro
12	85651	0	193	NÃO	SIM	NÃO	NÃO	NÃO	SIM	NÃO	Inclassificável
22	85815	0	102	NÃO	SIM	NÃO	NÃO	NÃO	SIM	IC	Inclassificável
71	86243	1	129	NÃO	NÃO	SIM	SIM	NÃO	NÃO	IC	Inclassificável
62	85650	1	167	NÃO	NÃO	SIM	SIM	NÃO	SIM	IC	Inclassificável
15	85884	0	99	NÃO	SIM	SIM	SIM	SIM	NÃO	IC	Inclassificável
26	86716	0	124	NÃO	SIM	SIM	SIM	NÃO	SIM	IC	Inclassificável
33	86801	0	177	SIM	SIM	NÃO	NÃO	SIM	SIM	IC	Inclassificável
52	85501	1	111	NÃO	SIM	SIM	SIM	NÃO	SIM	IC	Inclassificável
63	86798	1	167	SIM	SIM	NÃO	NÃO	SIM	SIM	IC	Inclassificável
27	85915	0	114	SIM	SIM	NÃO	SIM	SIM	SIM	IC	Inclassificável
31	86714	0	181	SIM	SIM	NÃO	SIM	SIM	SIM	IC	Inclassificável
51	85551	1	106	SIM	SIM	SIM	NÃO	SIM	SIM	IC	Inclassificável
60	85882	1	149	SIM	SIM	NÃO	SIM	SIM	SIM	IC	Inclassificável
11	85648	0	169	SIM	SIM	SIM	SIM	SIM	SIM	IC	acerto
18	85875	0	139	SIM	SIM	SIM	SIM	SIM	SIM	IC	acerto
49	86980	0	110	SIM	SIM	NÃO	SIM	SIM	SIM	SIM	acerto
53	85554	1	153	SIM	SIM	NÃO	SIM	SIM	SIM	SIM	acerto
54	85555	1	119	SIM	SIM	SIM	SIM	SIM	NÃO	SIM	acerto
64	85889	1	112	SIM	SIM	NÃO	SIM	SIM	SIM	SIM	acerto
67	85914	1	81	SIM	SIM	SIM	SIM	SIM	SIM	IC	acerto

66	85813	1	101	SIM	SIM	SIM	SIM	SIM	SIM	SIM	acerto
68	85870	1	152	SIM	SIM	SIM	SIM	SIM	SIM	SIM	acerto
70	85523	1	73	SIM	SIM	SIM	SIM	SIM	SIM	SIM	acerto
75	86258	1	176	SIM	SIM	SIM	SIM	SIM	SIM	SIM	acerto
77	86253	1	147	SIM	SIM	SIM	SIM	SIM	SIM	SIM	acerto
83	86330	1	190	SIM	SIM	SIM	SIM	SIM	SIM	SIM	acerto
88	86999	1	198	SIM	SIM	SIM	SIM	SIM	SIM	SIM	acerto

Classe 0: ausência de GVHD agudo; Classe 1: GVHD agudo; SIM : acerto do algoritmo; NÃO: erro do algoritmo IC: amostra inclassificável pelo algoritmo Bayesian Compound Covariate Predicto; Inclassificável : amostras não concordantes em 6 ou 7 métodos

ANEXO 5: Desempenho do classificador durante validação cruzada no grupo Teste (22 amostras)

Número Paciente	Array	Classe	Número de genes no classificador	Compound Covariate Predictor	Diagonal Linear Discriminant Analysis	1-Nearest Neighbor	1-Nearest Neighbor	Nearest Centroid	Support Vector Machines	Bayesian Compound Covariate Predictor	Classificação da amostra
4	85560	0	200	SIM	SIM	SIM	SIM	SIM	SIM	SIM	Acerto
5	85575	0	195	SIM	SIM	SIM	NÃO	SIM	SIM	SIM	Acerto
34	86313	0	187	SIM	SIM	SIM	SIM	SIM	SIM	SIM	Acerto
36	86259	0	203	SIM	SIM	SIM	SIM	SIM	SIM	SIM	Acerto
44	86323	0	195	SIM	SIM	SIM	SIM	SIM	SIM	SIM	Acerto
45	86324	0	204	SIM	SIM	SIM	SIM	SIM	SIM	SIM	Acerto
48	86984	0	213	SIM	SIM	SIM	SIM	SIM	SIM	SIM	Acerto
74	86312	1	219	SIM	SIM	SIM	SIM	SIM	SIM	SIM	Acerto
79	86790	1	211	SIM	SIM	SIM	NÃO	SIM	SIM	SIM	Acerto
82	86699	1	208	SIM	SIM	SIM	SIM	SIM	SIM	SIM	Acerto
87	86998	1	206	SIM	SIM	NÃO	SIM	SIM	SIM	SIM	Acerto
89	87000	1	219	SIM	SIM	SIM	SIM	SIM	SIM	SIM	Acerto
32	86319	0	202	NÃO	NÃO	NÃO	NÃO	NÃO	NÃO	NÃO	Erro
43	86325	0	184	NÃO	NÃO	NÃO	NÃO	NÃO	NÃO	NÃO	Erro
84	86804	1	198	NÃO	NÃO	NÃO	NÃO	NÃO	NÃO	NÃO	Erro
86	86997	1	203	NÃO	NÃO	NÃO	NÃO	NÃO	NÃO	IC	Erro
1	85553	0	202	NÃO	SIM	SIM	SIM	NÃO	SIM	IC	Inclassificável
20	85872	0	191	SIM	SIM	NÃO	SIM	SIM	NÃO	IC	Inclassificável
46	86981	0	225	NÃO	NÃO	NÃO	SIM	NÃO	SIM	IC	Inclassificável
50	85500	1	219	NÃO	SIM	NÃO	SIM	SIM	SIM	IC	Inclassificável
58	85581	1	217	SIM	SIM	NÃO	NÃO	SIM	NÃO	IC	Inclassificável
85	86982	1	204	NÃO	NÃO	SIM	SIM	NÃO	NÃO	IC	Inclassificável

Classe 0: ausência de GVHD agudo; Classe 1: GVHD agudo; SIM : acerto do algoritmo; NÃO: erro do algoritmo IC: amostra inclassificável pelo algoritmo Bayesian Compound Covariate Predictor; Inclassificável : amostras não concordantes em 6 ou 7 métodos

ANEXO 6. Genes diferencialmente expressos no grupo sem GVHD agudo(Classe 0) em relação ao grupo com GVHD agudo (classe 1)

Valor de p paramétrico	FDR*	Relação entre sem aGVHD/ a GVHD**	Clone	Símbolo do gene	Description
6.00E-07	0.00014	0.2512269	<u>H200000646</u>	<u>DEFA4</u>	defensin, alpha 4, corticostatin (DEFA4), mRNA.
1.30E-06	0.000151	0.4351157	<u>H200005875</u>	<u>RNASE3</u>	ribonuclease, RNase A family, 3 (eosinophil cationic protein) (RNASE3), mRNA.
5.30E-06	0.000412	0.3907423	<u>H200011085</u>	<u>CTSG</u>	cathepsin G (CTSG), mRNA.
2.56E-05	0.001272	0.4594606	<u>H200011001</u>	<u>ELA2</u>	elastase 2, neutrophil (ELA2), mRNA.
2.73E-05	0.001272	0.3871491	<u>H200000469</u>	<u>MPO</u>	myeloperoxidase (MPO), nuclear gene encoding mitochondrial protein, mRNA.
6.39E-05	0.002482	0.3997378	<u>H200006243</u>	<u>HPR</u>	haptoglobin-related protein (HPR), mRNA.
0.000152	0.004517	0.5441224	<u>H200004906</u>	<u>CDCA8</u>	cell division cycle associated 8 (CDCA8), mRNA.
0.0001551	0.004517	0.5544679	<u>H200000275</u>	<u>PSCD1</u>	pleckstrin homology, Sec7 and coiled-coil domains 1(cytohesin 1) (PSCD1), transcript variant 2
0.0002883	0.007464	0.4716303	<u>H200008642</u>	<u>MAOA</u>	monoamine oxidase A (MAOA), nuclear gene encoding mitochondrial protein, mRNA.
0.0003895	0.00825	0.5764731	<u>H200011744</u>	<u>ALOX15B</u>	arachidonate 15-lipoxygenase, type B (ALOX15B), transcript variant b, mRNA.
0.0003819	0.00825	0.4623782	<u>H200019509</u>	<u>BEX1</u>	brain expressed, X-linked 1 (BEX1), mRNA.
0.0005799	0.01126	1.9998148	<u>H200005341</u>	<u>HHAT</u>	hedgehog acyltransferase (HHAT), mRNA.
0.0007422	0.012352	1.7007395	<u>H200015643</u>	<u>PBEF1</u>	pre-B-cell colony enhancing factor 1 (PBEF1), mRNA.
0.0007035	0.012352	0.5340504	<u>H200014292</u>	<u>STK38L</u>	serine/threonine kinase 38 like (STK38L), mRNA.
0.0008171	0.012692	0.6156616	<u>H200009821</u>	<u>MYH10</u>	myosin, heavy chain 10, non-muscle (MYH10), mRNA.
0.0017139	0.023491	0.5751104	<u>H200014027</u>	<u>CA2</u>	carbonic anhydrase II (CA2), mRNA.
0.0020461	0.025092	0.6592785	<u>H200007941</u>	<u>TMOD1</u>	tropomodulin 1 (TMOD1), mRNA.
0.0028241	0.032901	0.6051023	<u>H200010342</u>	<u>BPI</u>	bactericidal/permeability-increasing protein (BPI), mRNA.
0.0042111	0.034887	0.6352583	<u>H200019041</u>	<u>ALAS2</u>	aminolevulinatase, delta-, synthase 2 (sideroblastic/hypochromic anemia) (ALAS2)
0.0046308	0.034887	0.4271345	<u>H200010278</u>	<u>BACH2</u>	BTB and CNC homology 1, basic leucine zipper transcription factor 2 (BACH2), mRNA.

0.0036813	0.034887	1.5869924	<u>H200019894</u>	<u>CACNA2D3</u>	calcium channel, voltage-dependent, alpha 2/delta 3 subunit (CACNA2D3), mRNA.
0.00388	0.034887	1.5087043	<u>H200005476</u>	<u>F3</u>	coagulation factor III (thromboplastin, tissue factor) (F3), mRNA.
0.0041758	0.034887	0.492066	<u>H200004352</u>	<u>HBD</u>	hemoglobin, delta (HBD), mRNA.
0.0032213	0.034887	0.6882021	<u>H200006961</u>	<u>HMBS</u>	hydroxymethylbilane synthase (HMBS), transcript variant 2, mRNA.
0.0046416	0.034887	0.6637731	<u>H200016970</u>	<u>LOC647979</u>	PREDICTED: hypothetical LOC647979, transcript variant 2 (LOC647979), mRNA.
0.0045961	0.034887	1.6945807	<u>H200011257</u>	<u>RIPK2</u>	receptor-interacting serine-threonine kinase 2 (RIPK2), mRNA.
0.0040391	0.034887	0.5521378	<u>H200014448</u>	<u>TMEM176B</u>	transmembrane protein 176B (TMEM176B), mRNA.
0.0043601	0.034887	0.5891223	<u>H200014375</u>	<u>ZNF524</u>	zinc finger protein 524 (ZNF524), mRNA.
0.0048386	0.035231	1.79034	<u>H200014737</u>	<u>PTGS2</u>	prostaglandin-endoperoxide synthase 2 (prostaglandin G/H synthase and cyclooxygenase) (PTGS2)
0.004996	0.035275	0.5955574	<u>H200009355</u>	<u>METTL7A</u>	methyltransferase like 7A (METTL7A), mRNA.
0.0060542	0.041489	0.5937465	<u>H200009424</u>	<u>FLJ20920</u>	hypothetical protein FLJ20920 (FLJ20920), mRNA.
0.0065685	0.043003	0.5124285	<u>H200012291</u>	<u>C21orf55</u>	chromosome 21 open reading frame 55 (C21orf55), transcript variant 2, mRNA.
0.0066442	0.043003	0.4954714	<u>H200015022</u>	<u>ITIH5</u>	inter-alpha (globulin) inhibitor H5 (ITIH5), transcript variant 2, mRNA.
0.0073688	0.046404	0.6501841	<u>H200015984</u>	<u>HBQ1</u>	hemoglobin, theta 1 (HBQ1), mRNA.
0.0079272	0.048606	1.8112116	<u>H200017533</u>	<u>THG1L</u>	TRNA-histidine guanylyltransferase 1-like (S. cerevisiae)
0.0093948	0.049021	0.6031507	<u>H200000316</u>	<u>ADH4</u>	alcohol dehydrogenase 4 (class II), pi polypeptide (ADH4), mRNA.
0.0088361	0.049021	1.6668666	<u>H200002660</u>	<u>C11orf1</u>	Chromosome 11 open reading frame 1
0.0100987	0.049021	0.6818692	<u>H200017053</u>	<u>ERAF</u>	erythroid associated factor (ERAF), mRNA.
0.0092834	0.049021	0.6138128	<u>H200010352</u>	<u>ETV1</u>	Ets variant gene 1
0.0100336	0.049021	0.5838595	<u>H200009720</u>	<u>GDF15</u>	growth differentiation factor 15 (GDF15), mRNA.
0.0096911	0.049021	1.7262492	<u>H200006773</u>	<u>NR4A3</u>	nuclear receptor subfamily 4, group A, member 3 (NR4A3), transcript variant 3, mRNA.
0.0088043	0.049021	0.6350028	<u>H200017397</u>	<u>PURG</u>	purine-rich element binding protein G (PURG), transcript variant A, mRNA.

0.0086433	0.049021	0.6896484	H200014442	TRIM65	tripartite motif-containing 65 (TRIM65), mRNA.
0.0110472	0.052531	0.5577647	H200005086	CRISP3	cysteine-rich secretory protein 3 (CRISP3), mRNA.
0.011398	0.053115	0.6403739	H200012248	ADAMTS2	ADAM metalloproteinase with thrombospondin type 1 motif, 2 (ADAMTS2), transcript variant 1, mRNA.
0.0121896	0.05569	0.5066405	H200000684	POU6F1	POU domain, class 6, transcription factor 1 (POU6F1), mRNA.
0.0134133	0.058968	0.5292769	H200017325	IFI27	Interferon, alpha-inducible protein 27
0.0144121	0.062186	0.5193429	H200014086	CHST10	Hypothetical protein MGC13219
0.015785	0.065677	1.3705017	H200009102	FLJ21272	Hypothetical protein FLJ21272
0.0155092	0.065677	1.5711061	H200007119	KLF5	Kruppel-like factor 5 (intestinal) (KLF5), mRNA.
0.0165361	0.067571	0.5766418	H200019974	RAD51	RAD51 homolog (RecA homolog, E. coli) (S. cerevisiae) (RAD51), transcript variant 2, mRNA.
0.0178731	0.070584	0.5798389	H200014134	FAM53B	family with sequence similarity 53, member B (FAM53B), mRNA.
0.0182561	0.070895	0.4410722	H200000507	SPTA1	Spectrin, alpha, erythrocytic 1 (elliptocytosis 2)
0.0210324	0.079969	1.4952658	H200013038	FBXL4	F-box and leucine-rich repeat protein 4
0.0212794	0.079969	1.8618499	H200000451	IL1A	interleukin 1, alpha (IL1A), mRNA.
0.0233858	0.08649	0.6256161	H200013469	E2F7	E2F transcription factor 7 (E2F7), mRNA.
0.0243609	0.088369	1.5952568	H200012604	IL1B	interleukin 1, beta (IL1B), mRNA.
0.0246524	0.088369	0.6415978	H200006346	SNCA	synuclein, alpha (non A4 component of amyloid precursor) (SNCA), transcript variant NACP112, mRNA.
0.0264695	0.093017	1.5323814	H200004325	ALPK3	alpha-kinase 3 (ALPK3), mRNA.
0.0267475	0.093017	0.5032782	H200016920	CASR	calcium-sensing receptor (hypocalciuric hypercalcemia 1, severe neonatal hyperparathyroidism) (CASR)
0.0272915	0.093514	0.700673	H200016982	AGRN	agrin (AGRN), mRNA.
0.0286879	0.094145	0.7408758	H200014514	ADARB2	adenosine deaminase, RNA-specific, B2 (RED2 homolog rat) (ADARB2), mRNA.
0.0280162	0.094145	1.5239038	H200013727	NDUFS7	NADH dehydrogenase (ubiquinone) Fe-S protein 7, 20kDa (NADH-coenzyme Q reductase)
0.0306808	0.096945	0.6991681	H200016988	ANAPC4	Anaphase promoting complex subunit 4
0.0304653	0.096945	0.659564	H200015094	ATXN7L1	Ataxin 7-like 1
0.032811	0.096945	0.6568402	H200011322	FLJ10292	mago-nashi homolog 2 (FLJ10292), mRNA.
0.0328653	0.096945	0.6508655	H200009918	IGHG1	Immunoglobulin heavy constant gamma 1 (G1m marker)

0.0316635	0.096945	0.6242003	<u>H200020117</u>	<u>MTMR1</u>	myotubularin related protein 1 (MTMR1), mRNA.
0.0307929	0.096945	0.6991987	<u>H200006991</u>	<u>TFF3</u>	trefoil factor 3 (intestinal) (TFF3), mRNA.

* FDR : False Discovery Ratio ou taxa de descobertas falsas

** Média geométrica da razão no grupo sem GVHD/ Média geométrica da razão expressão no grupo com GVHD agudo

ANEXO 7 – Carta de aprovação do Comitê de Ética em Pesquisa



São Paulo, 10 de fevereiro de 2006.
CEP 0101/06

Ilmo(a). Sr(a).

Pesquisador(a) ADRIANO DE MORAES ARANTES

Co-Investigadores: José Salvador Rodrigues de Oliveira; Maria Gerbase de Lima; Andrey Morgun; Natalia Shulzenko; Amador Gonçalves Primo Júnior

Disciplina/Departamento: Medicina de Urgência/Medicina da Universidade Federal de São Paulo/Hospital São Paulo

Patrocinador: CAPES.

CARTA DE APROVAÇÃO E PARECER CONSUBSTANCIADO DO COMITÊ DE ÉTICA INSTITUCIONAL

Ref: Projeto de pesquisa intitulado: “**Análise do perfil de expressão gênica em pacientes submetidos a transplante alogênico de células tronco hematopoéticas.**”.

CARACTERÍSTICA PRINCIPAL DO ESTUDO: Estudo genético em pacientes submetidos a transplante de células tronco hematopoéticas.

RISCOS ADICIONAIS PARA O PACIENTE: risco mínimo, desconforto leve com coleta de sangue.

OBJETIVOS: Analisar, através de microarray e PCR em tempo real, amostras de sangue periférico de pacientes submetidos a transplante alogênico de células tronco hematopoéticas.

RESUMO: Serão utilizadas amostras de sangue periférico de pacientes submetidos a transplante de medula óssea do Hospital São Paulo (Unifesp). Estas amostras serão obtidas no período pré-transplante, no dia da enxertia medular e 3 meses pós-transplante, além de amostras adicionais durante os seguintes eventos clínicos: doença do enxerto contra o hospedeiro aguda, doença do enxerto contra hospedeiro crônica, infecção e recaída da doença de base. Os dados clínicos dos pacientes serão retirados dos arquivos da Unidade de Transplantes. Serão utilizadas amostras de cerca de 100 pacientes, número estimado como sendo necessário para obtenção de resultados significantes nos estudos de microarray. Na fase de doação de células tronco hematopoéticas será coletada do doador alíquota de 5 ml diretamente do produto de aférese ou do produto de coleta de medula óssea para extração de RNA, e armazenadas a -80°C no Laboratório de Imunogenética da Unifesp até o momento do uso. Os padrões de expressão gênica serão analisados a partir de RNAs amplificados de células mononucleares do sangue periférico, por microarray. Para cada evento clínico, serão analisadas as diferenças entre os pacientes que evoluem mal e os que têm boa evolução clínica. Através desta análise, espera-se definir possíveis genes capazes de distinguir e prever as diferentes evoluções dos eventos clínicos citados.

FUNDAMENTOS E RACIONAL: Através desta análise, espera-se definir possíveis genes capazes de distinguir e prever as diferentes evoluções dos eventos clínicos citados.

MATERIAL E MÉTODO: descritos os procedimentos laboratoriais que serão realizados por equipe especializada.

TCLE: adequado de acordo com a Res 196/96.

OBJETIVO ACADÊMICO: doutorado.

ENTREGA DE RELATÓRIOS PARCIAIS AO CEP PREVISTOS PARA: 9/8/2006 e 5/2/2007.

O Comitê de Ética em Pesquisa da Universidade Federal de São Paulo/Hospital São Paulo **ANALISOU** e **APROVOU** o projeto de pesquisa referenciado.

1. Comunicar toda e qualquer alteração do projeto e termo de consentimento livre e esclarecido. Nestas circunstâncias a inclusão de pacientes deve ser temporariamente interrompida até a resposta do Comitê, após análise das mudanças propostas.
2. Comunicar imediatamente ao Comitê qualquer evento adverso ocorrido durante o desenvolvimento do estudo.
3. Os dados individuais de todas as etapas da pesquisa devem ser mantidos em local seguro por 5 anos para possível auditoria dos órgãos competentes.

Atenciosamente,



Prof. Dr. José Osmar Medina Pestana
Coordenador do Comitê de Ética em Pesquisa da
Universidade Federal de São Paulo/ Hospital São Paulo

ANEXO 8 – Termo de Consentimento Livre e Esclarecido

TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO

1. Título do projeto:

Análise do perfil de expressão gênica em pacientes submetidos a transplante de células tronco hematopoéticas .

2. Resumo do estudo :

Transplante alogênico de células tronco hematopoéticas é modalidade de tratamento que promove eficácia terapêutica para pacientes com doenças hematopoéticas como leucemias agudas, leucemias crônicas, mieloma, linfomas e depressões medulares. Para muitos pacientes, esta modalidade de tratamento permanece como única opção, oferecendo um potencial de cura ou controle a longo prazo. Uma complicação importante do transplante é a doença do enxerto contra o hospedeiro, quando as células infundidas do doador, em especial os linfócitos T, atacam tecidos do receptor provocando lesões de pele, diarreias e alterações de enzimas hepáticas. Nosso estudo, utilizando novas técnicas de estudo do genoma humano e de expressão gênica pretende abranger vários níveis biológicos de análise de doença do enxerto, visando identificar os genes que determinam a susceptibilidade, evolução e resposta ao tratamento.

3. Objetivo :

Você está sendo convidado a participar de uma pesquisa cujo objetivo é identificar fatores genéticos responsáveis pela susceptibilidade à doença do enxerto contra o hospedeiro que possam influenciar a evolução e resposta ao tratamento.

4. Procedimento:

Serão coletados 10 (dez) ml de sangue por venopunção periférica que serão armazenados em local adequado no laboratório de Imunogenética. O material será utilizado no projeto descrito acima, a ser realizado, conjuntamente, na disciplina de Hematologia e Hemoterapia e no setor de Imunogenética da UNIFESP, O material será armazenado, e poderá ser utilizado em projetos futuros, após aprovação destes pelo Comitê de Ética em Pesquisa da UNIFESP.

5. Descrição dos possíveis desconfortos:

Os únicos riscos existentes neste estudo são os associados à coleta de sangue, como dor, hematoma e sangramento no local da punção.

6. Benefícios da participação:

Não é esperado que os resultados destes testes tragam benefícios diretos a você ou que modifiquem o curso do seu tratamento. Entre os benefícios indiretos está o possível avanço do conhecimento médico, tornando possível a descoberta de tratamentos mais eficazes para a doença do enxerto. Além disto, um possível benefício futuro, seria a contribuição na descoberta de novos fatores genéticos que influenciam a evolução e a resposta ao tratamento atualmente utilizado.

7. Garantia de acesso:

Em qualquer etapa do estudo, você terá acesso aos profissionais responsáveis pela pesquisa para esclarecimento de eventuais dúvidas. O pesquisador responsável por este projeto é o médico Adriano de Moraes Arantes. As amostras ficarão armazenadas no Laboratório de Imunogenética da UNIFESP, à Rua Loefgreen, número 1235, telefone 2101-7100, sob responsabilidade da diretora do mesmo, Dra. Maria Gerbase de-Lima. Se você tiver alguma consideração ou dúvida sobre a ética da pesquisa, entre em contato com o Comitê de Ética em Pesquisa (CEP), à Rua Botucatu, 572, 1º andar, cj. 14, telefone: 5571-1062, fax: 5539-7162, e-mail: cepunifesp@epm.br.

8.Saída do estudo:

É garantida a liberdade de retirada do consentimento a qualquer momento e deixar de participar do estudo, sem qualquer prejuízo à continuidade de seu tratamento na instituição. Você continuará sendo o proprietário destas amostras de sangue e sempre terá o direito de determinar a destruição delas, a qualquer momento, bastando para isso, entrar em contato com o médico do estudo.

9.Confidencialidade:

Ao assinar este termo de consentimento livre, você estará concedendo aos seus médicos e aos membros envolvidos nesse estudo, acesso a seus registros médicos. Para protegê-lo contra o risco de perda de confidencialidade você não será identificado em publicações ou relatórios derivados deste estudo

10.Atualização:

Todos os pacientes terão direito de se manter atualizados sobre os resultados parciais da pesquisa que sejam de conhecimento dos pesquisadores.

11.Despesas e compensações:

Não há despesas pessoais aos participantes em qualquer fase do estudo, incluindo exames e consultas. Também não há compensação financeira relacionada à sua participação. Se houver qualquer despesa adicional ela será absorvida pelo orçamento da pesquisa. Aos participantes que necessitarem, será fornecido atestado médico para o período das consultas e coletas de exames.

12.Danos pessoais:

Não são previstos efeitos colaterais ou quaisquer danos pessoais associados ao estudo. Se houver qualquer dano pessoal diretamente causado pelo procedimento proposto no estudo, o participante terá direito a tratamento médico na instituição, bem como às indenizações legalmente estabelecidas.

13. Compromisso do pesquisador:

O pesquisador estabelece o compromisso de utilizar os dados e materiais coletados apenas para essa pesquisa.

Os propósitos do estudo, os procedimentos a serem realizados, seus desconfortos e riscos, as garantias de confidencialidade e de esclarecimentos permanentes ficaram perfeitamente claros para mim. Discuti com o médico Adriano de Moraes Arantes sobre a minha decisão em participar desse estudo. Também ficou claro que minha participação é isenta de despesas e que tenho garantia de acesso a tratamento hospitalar quando necessário. Concordo voluntariamente em participar deste estudo e poderei retirar meu consentimento a qualquer momento, antes ou durante o mesmo, sem penalidades, prejuízo ou perda de qualquer benefício que eu possa ter adquirido e sem comprometimento do meu atendimento neste serviço.

_____ data: __/__/____
Assinatura do paciente/ representante legal

_____ data: __/__/____
Assinatura da testemunha

Declaro que obtive de forma apropriada e voluntária o Consentimento Livre e Esclarecido deste paciente ou representante legal para a participação neste estudo.

_____ data: __/__/____
Adriano de Moraes Arantes
Pesquisador

8. Referências

1. Rizzo JD, Wingard JR, Tichelli A, Lee SJ, Van Lint MT, Burns LJ, Davies SM, Ferrara JL, Socié G. Recommended screening and preventive practices for long-term survivors after hematopoietic cell transplantation: joint recommendations of the European Group for Blood and Marrow Transplantation, Center for International Blood and Marrow Transplant Research, and the American Society for Blood and Marrow Transplantation (EBMT/CIBMTR/ASBMT). *Bone Marrow Transplant*. 2006 Feb;37(3):249-61.
2. Hahn T, McCarthy PL Jr, Zhang MJ, Wang D, Arora M, Frangoul H, Gale RP, Hale GA, Horan J, Isola L, Maziarz RT, van Rood JJ, Gupta V, Halter J, Reddy V, Tiberghien P, Litzow M, Anasetti C, Pavletic S, Ringdén O. Risk factors for acute graft-versus-host disease after human leukocyte antigen-identical sibling transplants for adults with leukemia *J Clin Oncol*. 2008 Dec 10;26(35):5728-34.
3. Arai S, Vogelsang GB. Management of graft-versus-host disease. *Blood Rev*. 2000 Dec;14(4):190-204
4. Przepiorka D, Weisdorf D, Martin P, Klingemann HG, Beatty P, Hows J, Thomas ED. 1994 Consensus Conference on Acute GVHD Grading . *Bone Marrow Transplant*. 1995 Jun;15(6):825-8.
5. Hill GR, Ferrara JL. The primacy of the gastrointestinal tract as a target organ of acute graft-versus-host disease: rationale for the use of cytokine shields in allogeneic bone marrow transplantation. *Blood*. 2000 May 1;95(9):2754-9.
6. Sorrow ML, Maris MB, Storer B, Sandmaier BM, Diaconescu R, Flowers C, Maloney DG, Storb R. Comparing morbidity and mortality of HLA-matched unrelated donor hematopoietic cell transplantation after nonmyeloablative and myeloablative conditioning: influence of pretransplantation comorbidities. *Blood*. 2004 Aug 15;104(4):961-8.
7. Mielcarek M, Storer BE, Sandmaier BM, Sorrow ML, Maloney DG, Petersdorf E, Martin PJ, Storb R. Comparable outcomes after nonmyeloablative hematopoietic cell transplantation with unrelated and related donors. *Biol Blood Marrow Transplant*. 2007 ; 13(12):1499-507.

8. Glucksberg H, Storb R, Fefer A, Buckner CD, Neiman PE, Clift RA, Lerner KG, Thomas ED. Clinical manifestations of graft-versus-host disease in human recipients of marrow from HL-A-matched sibling donors. *Transplantation*. 1974 ; 18(4):295-304.
9. Deeg HJ. How I treat refractory acute GVHD. *Blood*. 2007 May 15;109(10):4119-26.
10. Arai S, Margolis J, Zahurak M, Anders V, Vogelsang GB. Poor outcome in steroid-refractory graft-versus-host disease with antithymocyte globulin treatment. *Biol Blood Marrow Transplant*. 2002; 8(3):155-60.
11. Gale RP, Bortin MM, van Bekkum DW, Biggs JC, Dicke KA, Gluckman E, Good RA, Hoffmann RG, Kay HE, Kersey JH. Risk factors for acute graft-versus-host disease. *Br J Haematol*. 1987; 67(4):397-406.
12. Przepiora D, Shapiro S, Schwinghammer TL, Bloom EJ, Rosenfeld CS, Shadduck RK, Venkataramanan R. Cyclosporine and methylprednisolone after allogeneic marrow transplantation: association between low cyclosporine concentration and risk of acute graft-versus-host disease. *Bone Marrow Transplant*. 1991;7(6):461-5.
13. Eisner MD, August CS: Impact of donor and recipient characteristics on the development of acute and chronic graft-versus-host disease following pediatric bone marrow transplantation. *Bone Marrow Transplant*. 1995; 15:663-668.
14. Weisdorf D, Hakke R, Blazar B, Miller W, McGlave P, Ramsay N, Kersey J, Filipovich A. Risk factors for acute graft-versus-host disease in histocompatible donor bone marrow transplantation. *Transplantation*. 1991 ;51(6):1197-203.
15. Anasetti C, Beatty PG, Storb R, Martin PJ, Mori M, Sanders JE, Thomas ED, Hansen JA. Effect of HLA incompatibility on graft-versus-host disease, relapse, and survival after marrow transplantation for patients with leukemia or lymphoma. *Hum Immunol*. 1990 29(2):79-91.
16. Clave E, Rocha V, Talvensaari K, Busson M, Douay C, Appert ML, Rabian C, Carmagnat M, Garnier F, Filion A, Socié G, Gluckman E, Charron D, Toubert A Prognostic value of pretransplantation host thymic function in HLA-identical sibling hematopoietic stem cell transplantation. *Blood*. 2005; 15;105(6):2608-13.
17. Douek DC, McFarland RD, Keiser PH, Gage EA, Massey JM, Haynes BF, Polis MA, Haase AT, Feinberg MB, Sullivan JL, Jamieson BD, Zack JA, Picker LJ, Koup RA.

- Changes in thymic function with age and during the treatment of HIV infection. *Nature*. 1998; 17;396(6712):690-5.
18. Naylor K, Li G, Vallejo AN, Lee WW, Koetz K, Bryl E, Witkowski J, Fulbright J, Weyand CM, Goronzy JJ. The influence of age on T cell generation and TCR diversity. *J Immunol*. 2005; 174,7446-52.
 19. Weinberg K, Blazar BR, Wagner JE, Agura E, Hill BJ, Smogorzewska M, Koup RA, Betts MR, Collins RH, Douek DC. Factors affecting thymic function after allogeneic hematopoietic stem cell transplantation. *Blood*. 2001; 97,1458-66.
 20. Mielcarek M, Martin PJ, Leisenring W, Flowers ME, Maloney DG, Sandmaier BM, Maris MB, Storb R. Graft-versus-host disease after nonmyeloablative versus conventional hematopoietic stem cell transplantation. *Blood*. 2003; 15;102(2):756-62.
 21. Couriel DR, Saliba RM, Giralt S, Khouri I, Andersson B, de Lima M, Hosing C, Anderlini P, Donato M, Cleary K, Gajewski J, Neumann J, Ippoliti C, Rondon G, Cohen A, Champlin R. Acute and chronic graft-versus-host disease after ablative and nonmyeloablative conditioning for allogeneic hematopoietic transplantation. *Biol Blood Marrow Transplant*. 2004;10(3):178-85.
 22. Pasquini MC, Wang Z et Schneider L. Current use and outcome of hematopoietic stem cell transplantation, part I –CIBMTR Summary Slides).
 23. Hahn T, McCarthy PL Jr, Zhang MJ, Wang D, Arora M, Frangoul H, Gale RP, Hale GA, Horan J, Isola L, Maziarz RT, van Rood JJ, Gupta V, Halter J, Reddy V, Tiberghien P, Litzow M, Anasetti C, Pavletic S, Ringdén O. Risk factors for acute graft-versus-host disease after human leukocyte antigen-identical sibling transplants for adults with leukemia. *J Clin Oncol*. 2008; 10;26(35):5728-34.
 24. DeRisi JL, Iyer VR, Brown PO. Exploring the metabolic and genetic control of gene expression on a genomic scale. *Science*. 1997; 24;278(5338):680-6.
 25. Schena M, Shalon D, Davis RW, Brown PO. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science*. 1995; 20;270(5235):467-70.
 26. Brown PO, Botstein D. Exploring the new world of the genome with DNA microarrays. *Nat Genet*. 1999;21(1 Suppl):33-7.
 27. Golub TR, Slonim DK, Tamayo P, Huard C, Gaasenbeek M, Mesirov JP, Coller H, Loh ML, Downing JR, Caligiuri MA, Bloomfield CD, Lander ES. Molecular classification of

- cancer: class discovery and class prediction by gene expression monitoring. *Science*. 1999;15;286(5439):531-7.
28. Alizadeh AA, Eisen MB, Davis RE, Ma C, Lossos IS, Rosenwald A, Boldrick JC, Sabet H, Tran T, Yu X, Powell JJ, Yang L, Marti GE, Moore T, Hudson J Jr, Lu L, Lewis DB, Tibshirani R, Sherlock G, Chan WC, Greiner TC, Weisenburger DD, Armitage JO, Warnke R, Levy R, Wilson W, Grever MR, Byrd JC, Botstein D, Brown PO, Staudt LM. Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature*. 2000;403(6769):503-11.
29. Shipp MA, Ross KN, Tamayo P, Weng AP, Kutok JL, Aguiar RC, Gaasenbeek M, Angelo M, Reich M, Pinkus GS, Ray TS, Koval MA, Last KW, Norton A, Lister TA, Mesirov J, Neuberg DS, Lander ES, Aster JC, Golub TR. Diffuse large B-cell lymphoma outcome prediction by gene-expression profiling and supervised machine learning. *Nat Med*. 2002;8(1):68-74.
30. Rosenwald A, Wright G, Chan WC, Connors JM, Campo E, Fisher RI, Gascoyne RD, Muller-Hermelink HK, Smeland EB, Giltnane JM, Hurt EM, Zhao H, Averett L, Yang L, Wilson WH, Jaffe ES, Simon R, Klausner RD, Powell J, Duffey PL, Longo DL, Greiner TC, Weisenburger DD, Sanger WG, Dave BJ, Lynch JC, Vose J, Armitage JO, Montserrat E, López-Guillermo A, Grogan TM, Miller TP, LeBlanc M, Ott G, Kvaloy S, Delabie J, Holte H, Krajci P, Stokke T, Staudt LM; Lymphoma/Leukemia Molecular Profiling Project. The use of molecular profiling to predict survival after chemotherapy for diffuse large-B-cell lymphoma. *N Engl J Med*. 2002; 20;346(25):1937-47.
31. Yeoh EJ, Ross ME, Shurtleff SA, Williams WK, Patel D, Mahfouz R, Behm FG, Raimondi SC, Relling MV, Patel A, Cheng C, Campana D, Wilkins D, Zhou X, Li J, Liu H, Pui CH, Evans WE, Naeve C, Wong L, Downing JR. Classification, subtype discovery, and prediction of outcome in pediatric acute lymphoblastic leukemia by gene expression profiling. *Cancer Cell*. 2002;1(2):133-43.
32. Bullinger L, Döhner K, Bair E, Fröhling S, Schlenk RF, Tibshirani R, Döhner H, Pollack JR. Use of gene-expression profiling to identify prognostic subclasses in adult acute myeloid leukemia. *N Engl J Med*. 2004;15;350(16):1605-16.

33. Sarwal M, Chua MS, Kambham N, Hsieh SC, Satterwhite T, Masek M, Salvatierra O Jr. Molecular heterogeneity in acute renal allograft rejection identified by DNA microarray profiling. *N Engl J Med.* 2003;10;349(2):125-38.
34. Flechner SM, Kurian SM, Head SR, Sharp SM, Whisenant TC, Zhang J, Chismar JD, Horvath S, Mondala T, Gilmartin T, Cook DJ, Kay SA, Walker JR, Salomon DR. Kidney transplant rejection and tissue injury by gene profiling of biopsies and peripheral blood lymphocytes. *Am J Transplant.* 2004; 4(9):1475-89.
35. Morgun A, Shulzhenko N, Perez-Diez A, Diniz RV, Sanson GF, Almeida DR, Matzinger P, Gerbase-DeLima M. Molecular profiling improves diagnoses of rejection and infection in transplanted organs. *Circ Res.* 2006; 23;98(12):e74-83
36. Dupuy A, Simon RM. Critical review of published microarray studies for cancer outcome and guidelines on statistical analysis and reporting. *J Natl Cancer Inst.* 2007; 17;99(2):147-57.
37. Guyon I, Weston J, Barnhill S, and Vapnik V. Gene selection for cancer classification using support vector machines. *Mach Learn.* 2002; 46: 389-422.
38. Simon R, Radmacher MD, Dobbin K, McShane LM. Pitfalls in the use of DNA microarray data for diagnostic and prognostic classification. *J Natl Cancer Inst.* 2003; 1;95(1):14-8
39. Baron C, Somogyi R, Greller LD, Rineau V, Wilkinson P, Cho CR, Cameron MJ, Kelvin DJ, Chagnon P, Roy DC, Busque L, Sékaly RP, Perreault C. Prediction of graft-versus-host disease in humans by donor gene-expression profiling. *PLoS Med.* 2007; 4(1):e23.
40. Rowe V, Banovic T, MacDonald KP, Kuns R, Don AL, Morris ES, Burman AC, Bofinger HM, Clouston AD, Hill GR. Host B cells produce IL-10 following TBI and attenuate acute GVHD after allogeneic bone marrow transplantation. *Blood.* 2006; 1;108(7):2485-92.
41. Storb R, Deeg HJ, Whitehead J, Appelbaum F, Beatty P, Bensinger W, Buckner CD, Clift R, Doney K, Farewell V. Methotrexate and cyclosporine compared with cyclosporine alone for prophylaxis of acute graft versus host disease after marrow transplantation for leukemia *N Engl J Med.* 1986; 314(12):729-35.

42. Barrett AJ, Savani BN. Stem cell transplantation with reduced-intensity conditioning regimens: a review of ten years experience with new transplant concepts and new therapeutic agents. *Leukemia*. 2006; 20(10):1661-72.
43. Schattenberg AV, Levenska TH. Differences between the different conditioning regimens for allogeneic stem cell transplantation. *Curr Opin Oncol*. 2006 ;18(6):667-70.
44. Burroughs L, Mielcarek M, Leisenring W, Sandmaier BM, Maloney DG, Baron F, Martin PJ, Flowers ME, Forman SJ, Chauncey TR, Bruno B, Storb R. Extending postgrafting cyclosporine decreases the risk of severe graft-versus-host disease after nonmyeloablative hematopoietic cell transplantation. *Transplantation*. 2006; 27;81(6):818-25.
45. Guidelines for preventing opportunistic infection among hematopoietic stem cell transplant recipients. Recommendation of CDC the Infectious Disease Society of America, and the American Society of Blood and Marrow Transplantation. 2000 Vol 49 N RR-10.
46. Shulman HM, Sullivan KM, Weiden PL, McDonald GB, Striker GE, Sale GE, Hackman R, Tsoi MS, Storb R, Thomas ED. Chronic graft-versus-host syndrome in man. A long-term clinicopathologic study of 20 Seattle patients. *Am J Med*. 1980; 69(2):204-17.
47. Lee SJ, Klein J, Haagenson M, Baxter-Lowe LA, Confer DL, Eapen M, Fernandez-Vina M, Flomenberg N, Horowitz M, Hurley CK, Noreen H, Oudshoorn M, Petersdorf E, Setterholm M, Spellman S, Weisdorf D, Williams TM, Anasetti C. High-resolution donor-recipient HLA matching contributes to the success of unrelated donor marrow transplantation. *Blood*. 2007 Dec 15;110(13):4576-83.
48. Transplant Essential Data CIBMTR 2008 – www.cibmtr.org/data
49. Wang E, Miller LD, Ohnmacht GA, Liu ET, Marincola FM. High-fidelity mRNA amplification for gene profiling. *Nat Biotechnol*. 2000; 18(4):457-9
50. Zhao H, Hastie T, Whitfield ML, Børresen-Dale AL, Jeffrey SS. Optimization and evaluation of T7 based RNA linear amplification protocols for cDNA microarray analysis. *BMC Genomics*. 2002; 3(1):31.
51. Quackenbush J. Microarray data normalization and transformation. *Nat Genet*. 2002; 32 Suppl:496-501.

-
52. Yang IV, Chen E, Hasseman JP, Liang W, Frank BC, Wang S, Sharov V, Saeed AI, White J, Li J, Lee NH, Yeatman TJ, Quackenbush J. Within the fold: assessing differential expression measures and reproducibility in microarray assays. *Genome Biol.* 2002; 24;3(11).
 53. Yang YH, Dudoit S, Luu P, Lin DM, Peng V, Ngai J, Speed TP. Normalization for cDNA microarray data: a robust composite method addressing single and multiple slide systematic variation. *Nucleic Acids Res.* 2002 Feb 15;30(4):e15.
 54. Klein JP, Rizzo JD, Zhang MJ, Keiding N. Statistical methods for the analysis and presentation of the results of bone marrow transplants. Part I: unadjusted analysis. *Bone Marrow Transplant.* 2001;28(10):909-15.
 55. Simon R, Radmacher MD, Dobbin K, McShane LM. Pitfalls in the use of DNA microarray data for diagnostic and prognostic classification. *J Natl Cancer Inst.* 2003; 95(1):14-8.
 56. Benjamini Y. and Hochberg Y. Controlling the False Discovery Rate: a Practical and Powerful Approach to Multiple Testing *Journal of the Royal Stat. Soc.* 57, 289-300 (1995).
 57. Su AI, Wiltshire T, Batalov S, Lapp H, Ching KA, Block D, Zhang J, Soden R, Hayakawa M, Kreiman G, Cooke MP, Walker JR, Hogenesch JB. A gene atlas of the mouse and human protein-encoding transcriptomes. *Proc Natl Acad Sci U S A.* 2004; 101(16):6062-7
 58. Saeed AI, Sharov V, White J, Li J, Liang W, Bhagabati N, Braisted J, Klapa M, Currier T, Thiagarajan M, Sturn A, Snuffin M, Rezantsev A, Popov D, Ryltsov A, Kostukovich E, Borisovsky I, Liu Z, Vinsavich A, Trush V, Quackenbush J. TM4: a free, open-source system for microarray data management and analysis. *Biotechniques.* 2003; 34(2):374-8
 59. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM, Sherlock G. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet.* 2000; 25(1):25-9).
 60. Weisdorf D, Haake R, Blazar B, Miller W, McGlave P, Ramsay N, Kersey J, Filipovich A. Treatment of moderate/severe acute graft-versus-host disease after allogeneic bone marrow transplantation: an analysis of clinical risk features and outcome. *Blood.* 1990; 75(4):1024-30.

61. Saliba RM, de Lima M, Giralt S, Andersson B, Khouri IF, Hosing C, Ghosh S, Neumann J, Hsu Y, De Jesus J, Qazilbash MH, Champlin RE, Couriel DR. Hyperacute GVHD: risk factors, outcomes, and clinical implications. *Blood*. 2007; 109(7):2751-8.
62. Khoury H, Kashyap A, Adkins DR, Brown RA, Miller G, Vij R, Westervelt P, Trinkaus K, Goodnough LT, Hayashi RJ, Parker P, Forman SJ, DiPersio JF. Treatment of steroid-resistant acute graft-versus-host disease with anti-thymocyte globulin. *Bone Marrow Transplant*. 2001; 27(10):1059-64.
63. Oppenheim JJ, Yang D. Alarmins: chemotactic activators of immune responses. *Curr Opin Immunol*. 2005; 17(4):359-65.
64. Raponi M, Belly RT, Karp JE, Lancet JE, Atkins D, Wang Y. Microarray analysis reveals genetic pathways modulated by tipifarnib in acute myeloid leukemia. *BMC Cancer*. 2004; 4:56.
65. Korngold, B. & Sprent, J. Lethal graft-versus-host disease after bone marrow transplantation across minor histocompatibility barriers in mice. Prevention by removing mature T cells from marrow. *J. Exp. Med.* 1978; 148, 1687–1698 .
66. Korngold, R. & Sprent, J. Features of T cells causing H-2-restricted lethal graft-vs.-host disease across minor histocompatibility barriers. *J. Exp. Med.* 1982; 155, 872–883.
67. Martin PJ, Schoch G, Fisher L, Byers V, Appelbaum FR, McDonald GB, Storb R, Hansen JA. A retrospective analysis of therapy for acute graft-versus-host disease: secondary treatment. *Blood*. 1991 Apr 15;77(8):1821-8.
68. Gratwohl A, Hermans J, Apperley J, Arcese W, Bacigalupo A, Bandini G, di Bartolomeo P, Boogaerts M, Bosi A, Carreras E, et al. Acute graft-versus-host disease: grade and outcome in patients with chronic myelogenous leukemia. Working Party Chronic Leukemia of the European Group for Blood and Marrow Transplantation. *Blood*. 1995; 86(2):813-8.
69. Pasquini MC Impact of graft-versus-host disease on survival. *Best Pract Res Clin Haematol*. 2008; 21(2):193-204)
70. Ratanatharathorn V, Nash RA, Przepiorka D, Devine SM, Klein JL, Weisdorf D, Fay JW, Nademanee A, Antin JH, Christiansen NP, van der Jagt R, Herzig RH, Litzow MR, Wolff SN, Longo WL, Petersen FB, Karanes C, Avalos B, Storb R, Buell DN, Maher RM, Fitzsimmons WE, Wingard JR. Phase III study comparing methotrexate and tacrolimus

- (prograf, FK506) with methotrexate and cyclosporine for graft-versus-host disease prophylaxis after HLA-identical sibling bone marrow transplantation. *Blood*. 1998; 92(7):2303-14.
71. Kanda Y, Izutsu K, Hirai H, Sakamaki H, Iseki T, Kodera Y, Okamoto S, Mitsui H, Iwato K, Hirabayashi N, Furukawa T, Maruta A, Kasai M, Atsuta Y, Hamajima N, Hiraoka A, Kawa K. Effect of graft-versus-host disease on the outcome of bone marrow transplantation from an HLA-identical sibling donor using GVHD prophylaxis with cyclosporin A and methotrexate. *Leukemia*. 2004; 18(5):1013-9.
72. Cahn JY, Klein JP, Lee SJ, Milpied N, Blaise D, Antin JH, Leblond V, Ifrah N, Jouet JP, Loberiza F, Ringden O, Barrett AJ, Horowitz MM, Socié G; Société Française de Greffe de Moëlle et Thérapie Cellulaire; Dana Farber Cancer Institute; International Bone Marrow Transplant Registry. Prospective evaluation of 2 acute graft-versus-host (GVHD) grading systems: a joint Société Française de Greffe de Moëlle et Thérapie Cellulaire (SFGM-TC), Dana Farber Cancer Institute (DFCI), and International Bone Marrow Transplant Registry (IBMTR) prospective study. *Blood*. 2005; 106(4):1495-500.
73. Macmillan ML, Weisdorf DJ, Brunstein CG, Cao Q, DeFor TE, Verneris MR, Blazar BR, Wagner JE. Acute graft-versus-host disease after unrelated donor umbilical cord blood transplantation: Analysis of risk factors. *Blood*. 2008 Nov 7. [Epub ahead of print]
74. Kataoka I, Kami M, Takahashi S, Kodera Y, Miyawaki S, Hirabayashi N, Okamoto S, Matsumoto N, Miyazaki Y, Morishita Y, Asai O, Maruta A, Yoshida T, Imamura M, Hamajima N, Matsuo K, Harada M, Mineishi S; Japan Society for Hematopoietic Cell Transplantation Working Party. Clinical impact of graft-versus-host disease against leukemias not in remission at the time of allogeneic hematopoietic stem cell transplantation from related donors. The Japan Society for Hematopoietic Cell Transplantation Working Party. *Bone Marrow Transplant*. 2004; 34(8):711-9.
75. Chakravarti A, Little P. Nature, nurture and human disease. *Nature*. 2003; 421:412– 414.
76. Perou CM, Sorlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA, Pollack JR, Ross DT, Johnsen H, Akslén LA, Fluge O, Pergamenschikov A, Williams C, Zhu SX, Lonning PE, Borresen-Dale AL, Brown PO, Botstein D. Molecular portraits of human breast tumours. *Nature*. 2000; 406:747–752/

77. Marchionni L, Wilson RF, Wolff AC, Marinopoulos S, Parmigiani G, Bass EB, Goodman SN. Systematic review: gene expression profiling assays in early-stage breast cancer. *Ann Intern Med.* 2008; 148:358–369.
78. Paczesny S, Krijanovski OI, Braun TM, Choi SW, Clouthier SG, Kuick R, Misek DE, Cooke KR, Kitko CL, Weyand A, Bickley D, Jones D, Whitfield J, Reddy P, Levine JE, Hanash SM, Ferrara JL. A biomarker panel for acute graft versus host disease. *Blood.* 2008 Oct 2. [Epub ahead of print]
79. Rimkus C, Friederichs J, Boulesteix AL, Theisen J, Mages J, Becker K, Nekarda H, Rosenberg R, Janssen KP, Siewert JR. Microarray-based prediction of tumor response to neoadjuvant radiochemotherapy of patients with locally advanced rectal cancer. *Clin Gastroenterol Hepatol.* 2008; 6(1):53-61.
80. Bianchi F, Nuciforo P, Vecchi M, Bernard L, Tizzoni L, Marchetti A, Buttitta F, Felicioni L, Nicassio F, Di Fiore PP. Survival prediction of stage I lung adenocarcinomas by expression of 10 genes. *J Clin Invest.* 2007;117(11):3436-44.
81. Nakamura H, Komatsu K, Ayaki M, Kawamoto S, Murakami M, Uoshima N, Yagi T, Hasegawa T, Yasumi M, Karasuno T, Teshima H, Hiraoka A, Masaoka T. Serum levels of soluble IL-2 receptor, IL-12, IL-18, and IFN-gamma in patients with acute graft-versus-host disease after allogeneic bone marrow transplantation. *J Allergy Clin Immunol.* 2000; 106(2):S45-50. .
82. Shaiegan M, Iravani M, Babae GR, Ghavamzadeh A. Effect of IL-18 and sIL2R on aGVHD occurrence after hematopoietic stem cell transplantation in some Iranian patients. *Transpl Immunol.* 2006; 15:223-227.
83. Visentainer JE, Lieber SR, Persoli LB, Vigorito AC, Aranha FJ, de Brito Eid KA, Oliveira GB, Miranda EC, de Souza CA. Serum cytokine levels and acute graft-versus-host disease after HLA-identical hematopoietic stem cell transplantation. *Exp Hematol.* 2003; 31:1044-1050.
84. Holler E, Kolb HJ, Möller A, Kempeni J, Liesenfeld S, Pechumer H, Lehmacher W, Ruckdeschel G, Gleixner B, Riedner C. Increased serum levels of tumor necrosis factor alpha precede major complications of bone marrow transplantation. *Blood.* 1990; 75:1011-1016).

85. Wright G, Tan B, Rosenwald A, Hurt EH, Wiestner A, Staudt LM. A gene expression-based method to diagnose clinically distinct subgroups of diffuse large B cell lymphoma. *Proc Natl Acad Sci U S A*. 2003; 100(17):9991-6.
86. Reddy V.; Iturraspe J.A.; Tzolas A.C.; Meier-Kriesche H.U.; Schold J.; Wingard J.R. Low dendritic cell count after allogeneic hematopoietic stem cell transplantation predicts relapse, death, and acute graft-versus-host disease. *Blood*. 2004; 103: 4330-4335.
87. Matzinger P. The danger model: a renewed sense of self. *Science*. 2002; 296(5566):301-5.
88. Clift RA, Buckner CD, Appelbaum FR, Bearman SI, Petersen FB, Fisher LD, Anasetti C, Beatty P, Bensinger WI, Doney K. Allogeneic marrow transplantation in patients with acute myeloid leukemia in first remission: a randomized trial of two irradiation regimens. *Blood*. 1990; 76(9):1867-71.
89. O. Ringden, Viral infections and graft-vs.-host disease. In: S.J. Burakoff, H.J. Deeg, J. Ferrara and K. Atkinson, Editors, *Graft-vs.-Host Disease*, Marcel Dekker, New York (1990), p. 467
90. Gabay JE, Almeida RP. Antibiotic peptides and serineprotease homologs in human polymorphonuclear leukocytes: defensins and azurocidin. *Curr Opin Immunol*. 1993; 5:97-102.
91. Ganz T, Selsted ME, Szklarek D, Harwig SS, Kaher K, Bainton DF, Lehrer RI. Defensins. Natural peptide antibiotics of human neutrophils. *J Clin Invest*. 1985; 76:1427-35.
92. Wilde CG, Griffith JE, Marra MN, Snable JL, Scott RW. Purification and characterization of human neutrophil peptide 4, a novel member of the defensin family. *J Biol Chem*. 1989; 264:11200-3.
93. Lourenzoni MR, Namba AM, Caseli L, Degreve L, Zaniquelli MED. Study of the interaction of human defensins with cell membrane models: relationships between structure and biological activity. *J Phys Chem B*. 2007; 111: 11318-29.
94. Yang D, Biragyn A, Hoover DM, Lubkowski J, Oppenheim JJ. Multiple roles of antimicrobial defensins, cathelicidins, and eosinophil-derived neurotoxin in host defense. *Annu Rev Immunol*. 2004; 22:181-215.
95. Chertov O, Michiel DF, Xu L, Wang JM, Tani K, Murphy WJ, Longo DL, Taub DD, Oppenheim JJ. Identification of defensin-1, defensin-2, and CAP37/azurocidin as T-cell

- chemoattractant proteins released from interleukin-8-stimulated neutrophils. *J Biol Chem.* 1996; 271(6):2935-40.
96. Yang D, Chertov O, Bykovskaia SN, Chen Q, Buffo MJ, Shogan J, Anderson M, Schröder JM, Wang JM, Howard OM, Oppenheim JJ. Beta-defensins: linking innate and adaptive immunity through dendritic and T cell CCR6. *Science.* 1999; 286(5439):525-8.
97. Yang D, Chen Q, Chertov O, Oppenheim JJ. Human neutrophil defensins selectively chemoattract naive T and immature dendritic cells. *J Leukoc Biol.* 2000; 68(1):9-14.
98. Travis J. Structure, function, and control of neutrophil proteinases. *Am J Med;* 1988; 84:37-42.
99. Kubes P, Smith R, Grisham MD, Granger DN. Neutrophil-mediated proteolysis. Differential roles for cathepsin G and elastase. *Inflammation.* 1993; 17:321-332.
100. Padrines M, Wolf M, Walz A, Baggiolini M. Interleukin-8 processing by neutrophil elastase, cathepsin G and proteinase-3. *FEBS Lett* 1994; 352:231-235.)
101. Chertov O, Ueda H, Xu LL, Tani K, Murphy WJ, Wang JM, Howard OM, Sayers TJ, Oppenheim JJ. Identification of human neutrophil-derived cathepsin G and azurocidin/CAP37 as chemoattractants for mononuclear cells and neutrophils. *J Exp Med.* 1997; 186(5):739-47.
102. Hase-Yamazaki T, Aoki Y. Stimulation of human lymphocytes by cathepsin G. *Cell Immunol.* 1995;160:24-32.
103. Yamazaki T, Aoki Y. Cathepsin G binds to human lymphocytes. *J Leukoc Biol* 1997; 61:73-79.
104. Baggiolini M, Loetscher P, Moser B. Interleukin-8 and the chemokine family. *Int J Immunopharmacol.* 1995; 17(2):103-8.
105. Diller R, Palmes D, Dietl KH, Senninger N, Winde G, Spiegel HU. Interleukin-6, interleukin-8, and interleukin-10 in kidney transplantation: improved risk strategy? *FEBS Lett.* 1994; 352(2):231-5.
106. Ugucioni M, Meliconi R, Nesci S, Lucarelli G, Ceska M, Gasbarrini G, Facchini A. Elevated interleukin-8 serum concentrations in beta-thalassemia and graft-versus-host disease. *Blood.* 1993; 81(9):2252-6.
107. Cooke K. Chemokines and Graft-vs.-Host-Disease. In: James L.M. Ferrara KRC, H.Joachim Deeg, ed. *Graft-vs-Host-Disease (ed Third): Marcel Dekker;* 2005:125-153.

108. La Rosa DF, Rahman AH, Turka LA. The innate immune system in allograft rejection and tolerance. *J Immunol.* 2007; 178(12):7503-9.
109. Lofti R, Lotze MT. Eosinophils induce DC maturation, regulation immunity. *J Leukoc Biol.* 2008; 83(3):456-460.
110. Jacobsen EA, Taranova AG, Lee NA, Lee JJ. Eosinophils: singularly destructive effector cells or purveyors immunoregulation? *J Allergy Clin Immunol.* 2007; 119(6):1313-20.
111. Daneshpouy M, Socie G, Lemann M, Rivet J, Gluckman E, Janin A. Activated eosinophils in upper gastrointestinal tract of patients with graft-versus-host disease. *Blood.* 2002; 99:3033-3040.
112. Per Venge and Jonas Byström. Molecules in focus Eosinophil cationic protein (ECP) *The International Journal of Biochemistry & Cell Biology.* 1998; 30(4): 433-437.
113. Hällgren R, Bohman SO, Fredens K. Activated eosinophil infiltration and deposits of eosinophil cationic protein in renal allograft rejection. *Nephron.* 1991;59:266-270.
114. Weir MR, Hall-Craggs M, Shen SY, Posner JN, Alongi SV, Dagher FJ, Sadler JH.. The prognostic value of the eosinophil in acute renal allograft rejection. *Transplantation.* 1986; 41:709-712.
115. Ben-Ari Z, Dhillon AP, Garwood L, Rolles K, Davidson B, Burroughs AK. Prognostic value of eosinophils for therapeutic response in severe acute hepatic allograft rejection. *Transplant Proc.* 1996; 28:3624-3628.
116. Paris F, Fuks Z, Kang A, Capodici P, Juan G, Ehleiter D, Haimovitz-Friedman A, Cordon-Cardo C, Kolesnick R. Endothelial apoptosis as the primary lesion initiating intestinal radiation damage in mice. *Science.* 2001; 293(5528):293-7.
117. T. Teshima and J.L. Ferrara, Understanding the alloresponse: new approaches to graft-versus-host disease prevention. *Semin. Hematol.* 2002; 39: 15-22.
118. Yang D and Oppenheim JJ *ARTHRITIS & RHEUMATISM.* 2004; 50(11) : 3401-3403.
119. Cowland JB, Borregaard N. The individual regulation of granule protein mRNA levels during neutrophil maturation explains the heterogeneity of neutrophil granules. *J Leukoc Biol.* 1999; 66(6):989-95.
120. Rondelli D, Andrews RG, Hansen JA, Ryncarz R, Faerber MA, Anasetti C. Alloantigen presenting function of normal human CD34+ hematopoietic cells *Blood.* 1996; 88(7):2619-75.

121. Ryncarz RE, Anasetti C. Expression of CD86 on human marrow CD34(+) cells identifies immunocompetent committed precursors of macrophages and dendritic cells. *Blood*. 1998; 91(10):3892-900.
122. Deeg HJ, Spitzer TR, Cottler-Fox M, Cahill R, Pickle LW. Conditioning-related toxicity and acute graft-versus-host disease in patients given methotrexate/cyclosporine prophylaxis. *Bone Marrow Transplant*. 1991;7:193.
123. Dumler JS, Beschorner WE, Farmer ER, Di Gennaro KA, Saral R, Santos GW. Endothelial-cell injury in cutaneous acute graft-versus-host disease. *Am J Pathol*. 1989; 135(6):1097-103.
124. Ruutu T, Barosi G, Benjamin RJ, Clark RE, George JN, Gratwohl A, Holler E, Iacobelli M, Kentouche K, Lämmle B, Moake JL, Richardson P, Socié G, Zeigler Z, Niederwieser D, Barbui T; European Group for Blood and Marrow Transplantation; European LeukemiaNet.. Diagnostic criteria for hematopoietic stem cell transplant-associated microangiopathy: results of a consensus process by an International Working Group. *Haematologica*. 2007; 92 : 95–100.

125. Cheuk DK, Wang P, Lee TL, Chiang AK, Ha SY, Lau YL, Chan GC. Risk factors and mortality predictors of hepatic veno-occlusive disease after pediatric hematopoietic stem cell transplantation. *Bone Marrow Transplant*. 2007; 40(10):935-44.
126. T.R. Spitzer, Engraftment syndrome following hematopoietic stem cell transplantation. *Bone Marrow Transplantation*. 2001; 27: 893–898.
127. Gorak E, Geller N, Srinivasan R, Espinoza-Delgado I, Donohue T, Barrett AJ, Suffredini A, Childs R. Engraftment syndrome after nonmyeloablative allogeneic hematopoietic stem cell transplantation: incidence and effects on survival. *Biol Blood Marrow Transplant*. 2005; 11(7):542-50.
128. Majhail NS, Parks K, Defor TE, Weisdorf DJ. Diffuse alveolar hemorrhage and infection-associated alveolar hemorrhage following hematopoietic stem cell transplantation: related and high-risk clinical syndromes. *Biol Blood Marrow Transplant*. 2006; 12(10):1038-46.

129. Woywodt A, Scheer J, Hambach L, Buchholz S, Ganser A, Haller H, Hertenstein B, Haubitz M. Circulating endothelial cells as a marker of endothelial damage in allogeneic hematopoietic stem cell transplantation. *Blood*. 2004; 103(9):3603-5.
130. Bahlmann FH, De Groot K, Spandau JM, Landry AL, Hertel B, Duckert T, Boehm SM, Menne J, Haller H, Fliser D. Erythropoietin regulates endothelial progenitor cells. *Blood*. 2004 Feb 1;103(3):921-6).
131. Der SD, Zhou A, Williams BR, Silverman RH. Identification of genes differentially regulated by interferon alpha, beta, or gamma using oligonucleotide arrays. *Proc Natl Acad Sci U S A*. 1998; 95(26):15623-8.
132. Stark GR, Kerr IM, Williams BRG, Silverman RH, Schreiber RD. How cells respond to interferons. *Annu Rev Biochem*. 1998; 67:227-264
133. Ichiba T, Teshima T, Kuick R, Misek DE, Liu C, Takada Y, Maeda Y, Reddy P, Williams DL, Hanash SM, Ferrara JL. Early changes in gene expression profiles of hepatic GVHD uncovered by oligonucleotide microarrays *Blood*. 2003; 102(2):763-71.
134. Teshima and J.L. Ferrara, Understanding the alloresponse: new approaches to graft-versus-host disease prevention. *Semin. Hematol*. 2002; 39: 15-22.
135. A. Velardi, P. Varese, A. Terenzi, C. Dembech, N. Albi, C.E. Grossi, L. Moretta, M.F. Martelli, F. Grignani and M.C. Mingari, Lymphokine production by T-cell clones after human bone marrow transplantation. *Blood*. 1989; 74:1665-1672.
136. H. Goker, I.C. Haznedaroglu and N.J. Chao, Acute graft-vs-host disease: pathobiology and management. *Exp. Hematol*. 2001; 29:259-277.
137. Baker MB, Altman NH, Podack ER, Levy RB. The role of cell-mediated cytotoxicity in acute GVHD after MHC-matched allogeneic bone marrow transplantation in mice. *J Exp Med*. 1996;183: 2645-2656.
138. Hattori K, Hirano T, Miyajima H. Differential effects of anti-Fas ligand and anti-tumor necrosis factor alpha antibodies on acute graft-versus-host disease pathologies. *Blood*. 1998; 91: 4051-4055
139. N . Takahashi , N . Sato , S . Takahashi , A . Tojo. Gene-expression profiles of peripheral blood mononuclear cell subpopulations in acute graft-vs-host disease following cord blood transplantation. *Experimental Hematology*. 2008; 36 (12) : 1760 - 1770.

140. Wysocki CA, Panoskaltsis-Mortari A, Blazar BR, Serody JS Leukocyte migration and graft-versus-host disease. : Blood. 2005; 105(11):4191-9.
141. Serody JS, Burkett SE, Panoskaltsis-Mortari A, Ng-Cashin J, McMahon E, Matsushima GK, Lira SA, Cook DN, Blazar BR. T-lymphocyte production of macrophage inflammatory protein-1alpha is critical to the recruitment of CD8(+) T cells to the liver, lung, and spleen during graft-versus-host disease. Blood. 2000;96: 2973-2980.
142. Wysocki CA, Burkett SB, Panoskaltsis-Mortari A, Kirby SL, Luster AD, McKinnon K, Blazar BR, Serody JS. Differential roles for CCR5 expression on donor T cells during graft-versus-host disease based on pretransplant conditioning. J Immunol. 2004;173: 845-854
143. Louvet C, Chiffolleau E, Heslan M, Tesson L, Heslan JM, Brion R, Bériou G, Guillonneau C, Khalife J, Anegon I, Cuturi MC. Identification of a new member of the CD20/Fcepsilon RIbeta family overexpressed in tolerated allografts. Am J Transplant. 2005 Sep;5(9):2143-53.
144. Dinarello CA. Biologic basis for interleukin-1 in disease. Blood. 1996; 87(6); 2095-147.
145. Inohara, N., del Peso, L., Koseki, T., Chen, S. & Nunez, G. RICK, a novel protein kinase containing a caspase recruitment domain, interacts with CLARP and regulates CD95-mediated apoptosis. *J. Biol. Chem.*1998; 273: 12296–12300.
146. McCarthy, J. V., Ni, J. & Dixit, V. M. RIP2 is a novel NF-κB-activating and cell death-inducing kinase. *J. Biol. Chem.*1998;273:16968–16975.
147. Kelliher MA, Grimm S, Ishida Y, Kuo F, Stanger BZ, Leder P .The death domain kinase RIP mediates the TNF-induced NF-kappaB signal. *Immunity*. 1998 Mar;8(3):297-303.
148. Kobayashi K, Inohara N, Hernandez LD, Galán JE, Núñez G, Janeway CA, Medzhitov R, Flavell RA. RICK/Rip2/CARDIAK mediates signalling for receptors of the innate and adaptive immune systems *Nature*. 2002 Mar 14; 416(6877):194-9).
149. Rowbottom AW, Riches PG, Downie C, Hobbs JR. Monitoring cytokine production in peripheral blood during acute graft-versus-host disease following allogeneic bone marrow transplantation. *Bone Marrow Transplant*. 1993;12:635-641.
150. Tanaka J, Imamura M, Kasai M, Sakurada K, Miyazaki T. Cytokine gene expression after allogeneic bone marrow transplantation. *Leuk Lymphoma*. 1995;16:413-418.

151. Tanaka J, Imamura M, Kasai M, et al. Cytokine gene expression in peripheral blood mononuclear cells during graft-versus-host disease after allogeneic bone marrow transplantation. *Br J Haematol.* 1993;85:558-565.
152. Carayol G, Bourhis JH, Guillard M, Bosq J, Paillet C, Castagna L, Vernant JP, Pico JL, Hayat M, Chouaib S, Caignard A. Quantitative analysis of T helper 1, T helper 2, and inflammatory cytokine expression in patients after allogeneic bone marrow transplantation: relationship with the occurrence of acute graft-versus-host disease. *Transplantation.* 1997;63(9):1307-13).
153. Liem LM, van Houwelingen HC, Goulmy E. Serum cytokine levels after HLA-identical bone marrow transplantation. *Transplantation.* 1998; 66:863-871).
154. S. Abhyankar, D.G. Gilliland and J.L.M. Ferrara, Interleukin 1 is a critical effector molecule during cytokine dysregulation in graft-versus-host disease to minor histocompatibility antigens. *Transplantation.* 1993; 56: 1518–1523.
155. G.R. Hill, T. Teshima, A. Gerbitz, L. Pan, K.R. Cooke, Y.S. Brinson, J.M. Crawford and J.L. Ferrara, Differential roles of IL-1 and TNF- α on graft-versus-host disease and graft versus leukemia. *J. Clin. Invest.* 1999;104 : 459–467
156. P.L. McCarthy, S. Abhyankar, S. Neben, G. Newman, C. Sieff, R.C. Thompson, S.J. Burakoff and J.L.M. Ferrara, Inhibition of interleukin-1 by an interleukin-1 receptor antagonist prevents graft-versus-host disease. *Blood.* 1991; 78:1915–1918.
157. J.H. Antin, D. Weisdorf, D. Neuberg, R. Nicklow, S. Clouthier, S.J. Lee, E. Alyea, C. McGarigle, B.R. Blazar, S. Sonis, R.J. Soiffer and J.L. Ferrara, Interleukin-1 blockade does not prevent acute graft-versus-host disease: results of a randomized, double-blind, placebo-controlled trial of interleukin-1 receptor antagonist in allogeneic bone marrow transplantation. *Blood.* 2002; 100:3479–3482).
158. Jost PJ, Ruland J. Aberrant NF-kappaB signaling in lymphoma: mechanisms, consequences, and therapeutic implications. *Blood.* 2007; 109(7):2700-7).
159. Ward C, Walker A, Dransfield I, Haslett C, Rossi AG. Regulation of granulocyte apoptosis by NF-kappaB. *Biochem Soc Trans.* 2004; 32(Pt3):465-7.
160. Jia SH, Li Y, Parodo J, Kapus A, Fan L, Rotstein OD, Marshall JC. Pre-B cell colony-enhancing factor inhibits neutrophil apoptosis in experimental inflammation and clinical sepsis. *J Clin Invest.* 2004; 113(9):1318-27.

161. McConnell BB, Ghaleb AM, Nandan MO, Yang VW. The diverse functions of Krüppel-like factors 4 and 5 in epithelial biology and pathobiology. *Bioessays*. 2007; 29(6):549-57.
162. Yang XO, Doty RT, Hicks JS, Willerford DM. Regulation of T-cell receptor D beta 1 promoter by KLF5 through reiterated GC-rich motifs. *Blood*. 2003; 101(11):4492-9.
163. Li QX, Ke N, Sundaram R, Wong-Staal F. NR4A1, 2, 3--an orphan nuclear hormone receptor family involved in cell apoptosis and carcinogenesis. *Histol Histopathol*. 2006; 21(5):533-40.
164. Sun K, Wilkins DE, Anver MR, Sayers TJ, Panoskaltsis-Mortari A, Blazar BR, Welniak LA, Murphy WJ. Differential effects of proteasome inhibition by bortezomib on murine acute graft-versus-host disease (GVHD): delayed administration of bortezomib results in increased GVHD-dependent gastrointestinal toxicity. *Blood*. 2005; 106(9):3293-9.

ABSTRACT

Background: Allogeneic hematopoietic stem cell transplantation is an important last resort therapy for hematological diseases. Unfortunately, the success of a large proportion of these transplants is limited by graft-versus-host disease (GVHD). Currently known risk factors for acute GVHD (histoincompatibility, sex mismatch, older patients, previous pregnancies) do not provide a precise estimate of individual patient risk and do not help for individualization of the therapy. Early identification of those patients who will develop aGVHD may allow for individualized treatment, and also for the reduction of unnecessary treatment for those patients not at risk. Nowadays, however, there is no diagnostic method that allows prediction of aGVHD.

Objectives: The goal of our study was to reveal a gene expression profile that would predict the occurrence of aGVHD. In addition, using enrichment of gene ontology categories, to analyze differentially expressed genes in order to better understand biology of the events preceding aGVHD.

Material and methods: we collected blood samples from 89 recipients of myeloablative and reduced conditioning regimen HLA-identical sibling allogeneic hematopoietic stem cell transplants at the time of successful engraftment. We isolated total RNA from the peripheral blood mononuclear cells, amplified it, labeled, and co-hybridized to the microarray slides containing probes for 22,000 genes. The patients were divided into training and test groups, the former used to build a model discriminating patients with and without aGVHD and the latter - to test the model on independent samples. We selected the informative genes using "recursive feature elimination" method followed by seven different multivariate classification algorithms in order to establish a molecular classifier in the training set. Then we validated this new classifier in the test set of patients. We found differentially expressed genes using T-test and accepted those with estimated false discovery rate below 10%. We have used Biobase Explain to find enrichment of functional groups among differentially expressed genes.

Results: We found a molecular classifier comprised by 233 gene probes in the training set of samples which were selected based on the most accurate classification. In the test group of samples, we found that 80% of patients could be classified based on the concordance between classification methods as described above. For these patients, the classifier showed 75% of a predictive accuracy (71% of sensitivity and 78% of specificity). Analysis of functional annotations of differentially expressed genes showed that patients that developed acute GVHD have increased expression of antimicrobial genes, hemaglobin metabolism genes and alarmins. In these samples we also observed decreased expression of IL-1 and other genes involved in NF- κ B activation. Several genes up-regulated before aGVHD were associated with multiple types of precursor cells.

Conclusion: Our results show that molecular profiling is able to identify patients under high risk of acute GVHD at the time of engraftment. Establishing of a predictive diagnostic method for aGVHD is the first step of individualization of therapeutic strategy after hematopoietic stem cell transplantation. In addition, the results of functional enrichment analysis and expression in different cell populations suggest that early events during engraftment involving precursor cell populations might predefine results of interaction between stem cell allograft and patient body.

Key words: stem cell transplantation, acute graft versus host disease, microarray, molecular profiling.

BIBLIOGRAFIA CONSULTADA

1. Rother ET, Braga MER. Como elaborar sua tese: estrutura e referências. 2ª edição, 2005, 122p.
2. International Committee of Medical Journal Editors. Uniform requirements for manuscripts submitted to biomedical journals. Writing and editing for biomedical publication (text on the internet). Vancouver (Canada); update 2004 Oct (cited 2005 Jan 21). Available from:
<http://www.icmje.org>