

UNIVERSIDADE FEDERAL DO RIO GRANDE - FURG  
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIAS FISIOLÓGICAS:  
FISIOLOGIA ANIMAL COMPARADA

**"Mutações no gene OCT4 em células eritroleucêmicas  
humanas: implicações sobre o fenótipo de resistência  
a múltiplas drogas (MDR)"**

***Bruno Rodrigues de Oliveira***

Dissertação a ser defendida no âmbito do  
Programa de Pós-Graduação em Ciências  
Fisiológicas – Fisiologia Animal Comparada,  
como parte dos requisitos para obtenção do  
título de MESTRE em Fisiologia Animal  
Comparada.

Orientador: Prof. Dr. Luís Fernando Marins

Coorientadora: Prof<sup>a</sup>. Dr<sup>a</sup>. Gilma Santos Trindade

**Rio Grande, 2012.**

## **Agradecimentos**

Agradeço inicialmente a Prof<sup>a</sup>. Dra. Vivian Mary Barral Dodd Rumjanek por ter aceitado compor a banca avaliadora de minha dissertação de mestrado, se dispondo a vir de longe para trazer suas contribuições. Pessoa muito querida, exemplo de profissional. Agradeço a Prof<sup>a</sup>. Dra. Ana Paula Horn, por também ter aceito compor a banca de meu trabalho, trazendo contribuições muito importantes.

É com muito carinho que agradeço a Gilma e a Daka por terem me dado a oportunidade de trabalhar com esse assunto que cada vez mais me apaixono e quero correr atrás ainda bastante, tudo graças a chance que me deram, a confiança que depositaram em mim desde o início. A Gilma é como uma mãe, sempre pronta para ajudar no que estiver ao seu alcance, com muita compreensão e carinho. A Daka, uma pessoa muito inteligente, que vê nas pessoas um lado humano com facilidade, interessada e sempre disposta também a me auxiliar, sem precisar receber nada em troca, adotei-a como minha irmã mais velha. Nunca esquecerei seu carinho. E grato também por toda ajuda e apoio moral dado pela professora Ana Paula Votto.

É claro que não poderia me esquecer do Luf, que aceitou o desafio de me orientar no mestrado, mesmo sabendo que eu nunca tinha trabalhado com biologia molecular. Sua dedicação as melhorias no trabalho são notórias e seu empenho em realizar algo de qualidade, com um foco na publicação é algo que pretendo aprender muito ainda. Também possui um lado preocupado com seus alunos que demonstra brandamente nos momentos de necessidade. Está sempre por perto, sempre disposto a auxiliar, realmente um orientador muito presente.

Grato pelos amigos que conquistei durante a pós-graduação (Bin, Guerreiro, André, Gustavo, Nana, Marília, Márcio Geihs, Figueiredo, dentre outros) e também aqueles que

sempre estiveram ao meu lado (Renan, Marcos, Werner, Marcel, Michel) e outros que o curso de Biologia me trouxe (Alexandre, Bruna, Carol Reyes, Peixoto, Maiara, Darlan ...).. Essas pessoas tornaram possível passar por dias difíceis, muitas vezes como é de costume em laboratório , experimentos que não deram certo, prazos estourando... mesmo assim no final do dia ou no final de semana, ria, ria alegremente com essas pessoas e com os mais próximos até dividia situações. Sem esquecer é claro da Karine, que tanto me acompanhou em finais de semana em dias até tarde na furg, aprendendo juntos, mas também muitas vezes apenas pela companhia, grato mesmo por tudo que passamos e pela relação que hoje temos.

Minha Família, sempre ao meu lado, servindo de apoio fundamental para meu caminho ser seguido sem fraquejar. Confesso que cheguei a pensar em desistir, quando nada parecia dar certo, quando o tempo já não estava a meu favor, mas nessas horas que tive o apoio para seguir a decisão que quisesse de meus pais, e certo de que teria apoio independente de minha escolha, encontrei forças para continuar.



## Sumário

|  |           |
|--|-----------|
| Resumo .....   | 04        |
| Palavras-chave .....   | 05        |
| Abstract .....   | 06        |
| Keywords .....   | 07        |
| Introdução .....   | 08        |
| Objetivo geral .....   | 13        |
| Objetivos específicos .....  | 13        |
| Material e métodos .....   | 14        |
| Resultados.....  | 21        |
| Discussão .....  | 28        |
| Conclusões.....  | 34        |
| Perspectivas.....  | 35        |
| Referências.....   | 36        |
| <br>   |           |
| Artigo em inglês: <b>OCT4 mutations in human erythroleukemic cells:<br/>implications for multiple drug resistance (MDR) phenotype.....</b> | <b>41</b> |

## Resumo

O fator de transcrição OCT4 é um importante marcador de células tronco e tem sido relacionado com o conceito de células tronco tumorais (CTTs). Recentemente, ele tem sido também relacionado ao fenótipo de resistência a múltiplas drogas (MDR). O objetivo deste trabalho foi testar se o OCT4 está ligado ao fenótipo MDR em células eritroleucêmicas derivadas da linhagem LMC-K562. Para isso, foi realizada a análise de expressão de genes associados à superfamília de transportadores ABC (*ATP Binding Cassette*) e, também, de fatores de transcrição relacionados a células tronco. Os primeiros resultados apontaram uma relação direta entre OCT4 e transportadores ABC na linhagem MDR derivada de K562 (Lucena). O sequenciamento de promotores ABC não revelou qualquer mutação que pudesse explicar a expressão diferenciada do OCT4 na linhagem MDR. Posteriormente, o sequenciamento da região contendo o domínio homeobox do gene OCT4 de ambas as linhagens celulares evidenciou, pela primeira vez, que este fator de transcrição é alvo de mutações que podem estar relacionados com o fenótipo MDR. As mutações encontradas implicam substituições de vários aminoácidos em ambas as linhagens celulares. K562 teve sete substituições (três exclusivas), enquanto Lucena teve 13 (nove exclusivas). Além disso, uma busca *in silico* por motivos de fosforilação dentro da sequência de aminoácidos comparada, revelou que o OCT4 humano normal possui sete motivos potenciais de fosforilação. Entretanto, K562 perdeu um motivo e Lucena dois deles. Moléculas com diferentes padrões de fosforilação podem ter sua função modificada. Estes resultados indicam o OCT4 com um alvo potencial para o tratamento do câncer, especialmente aqueles resistentes à quimioterapia.

**Palavras-chave:** K562, Lucena, expressão gênica, transportadores ABC, fatores de transcrição, motivos de fosforilação.

## **Abstract**

The OCT4 transcription factor is a crucial stem cells marker and it has been related to the cancer stem cells (CSCs) concept. More recently, it has also been associated to the multiple drug resistance (MDR) phenotype. Our objective was to test whether OCT4 is linked to the MDR phenotype in erythroleukemic cells derived from LMC-K562 line. For this, we performed expression analyses of genes associated to the ATP Binding Cassette (ABC) superfamily, as well as stem cells-related transcription factors. Our first results pointed out the straight correlation between OCT4 and ABC transporters in K562-derivative MDR (Lucena) cells. Sequencing of ABC promoters did not reveal any mutation that could explain the differential expression of OCT4 in Lucena cells. Furthermore, sequencing of the homeobox domain region from the OCT4 gene isolated from both cell lines evinced, for the first time, that this transcription factor is a target of mutations and might be related to the MDR phenotype. The encountered mutations implied in several amino acids substitutions in both cell lines. K562 had seven amino acids substituted (three of them exclusive), while Lucena had 13 substitutions (nine of them exclusive). In addition, an *in silico* search for phosphorylation motifs within the amino acid stretch compared showed that human normal OCT4 has seven potential phosphorylation motifs. However, K562 has lost one phosphorylation motif and Lucena two of them. Molecules presenting different phosphorylation patterns may have its function modified. These findings bring OCT4 as an important target for cancer treatment, especially those resistant to chemotherapy.



**Keywords:** K562, Lucena, gene expression, ABC transporters, transcription factors, phosphorylation motifs

## Introdução

O conhecimento a respeito da maioria dos tumores revela a presença de subpopulações de células tronco (CTs) em seu interior (Alison et al., 2012). Num tecido normal, essas células são responsáveis pelo processo de renovação celular devido ao seu grande potencial de expansão e autorrenovação, o que está associado à sua longa vida. As CTs apresentam três propriedades básicas. A primeira é a diferenciação, que é extremamente importante na formação e reposição de diferentes tecidos. A diferenciação das CTs ocorre de forma progressiva e dá origem a tipos celulares específicos em estrutura e função, porém de vida curta. A segunda, é a habilidade de autorrenovação, formando novas CTs com as mesmas características de suas células de origem, mantendo a capacidade de proliferação, expansão e diferenciação inalteradas, o que possibilita a manutenção do estoque de CTs. A terceira é a capacidade de controle homeostático, modulando e balanceando a diferenciação e autorrenovação de acordo com estímulos ambientais, bem como suas limitações genéticas (Dalerba et al., 2007).

Devido ao fato de as CTs serem as únicas células de vida longa presentes na maioria dos tecidos, é provável que estas tenham maior probabilidade de acúmulo das mutações necessárias para o desencadeamento do processo carcinogênico. As mutações acumuladas nessas células podem acarretar perda da capacidade de reparo do DNA e do processo de apoptose, contribuindo para a formação e sobrevivência tumoral (HU et al., 2008). Dentro desse contexto, surge o conceito de células tronco tumorais (CTTs). Os tumores são compostos por populações heterogêneas de células sustentados por CTTs, as quais também possuem a capacidade de diferenciação ou

plasticidade e autorrenovação. Além disso, as CTTs compartilham da capacidade das CTs de migração, adesão a diversos tecidos e elevada taxa de proliferação (KLONISCH et al., 2008). O modelo de CTTs foi, primeiro, descrito em células com Leucemia Mielóide Crônica (LMC), onde se encontra um pequeno grupo de marcadores de superfície celular com o padrão CD34+CD38-, típico de células embrionárias, levando à conclusão de que células leucêmicas possuem uma origem primitiva. Os eventos-chave para a transformação maligna das células leucêmicas ocorrem nas células tronco hematopoiéticas, originando as CTTs (Bonnet e Dick, 1997).

Diante dessas novas perspectivas, as CTTs passam a ser um alvo em potencial para o tratamento oncogênico. Os métodos atuais de combate aos variados tipos de câncer se atêm em diminuir a massa do tumor, porém não conseguem eliminar todas as células cancerígenas, especialmente as CTTs, as quais podem recriar a massa tumoral original. Isso ocorre, porque os quimioterápicos convencionais atuam nas fases do ciclo celular em que as células estão em intensa proliferação. Contudo, estudos revelam que, tanto CT normais, quanto CTTs, permanecem em G0, um estado quiescente sem muita atividade mitótica (Bertolini et al., 2009; Alison et al., 2012; Nguyen et al., 2012). Desse modo, as drogas antineoplásicas não surtem efeito nas CTTs graças a este microambiente diferenciado. E, por isso, uma quimioterapia desenhada especificamente para CTTs não atingiria rapidamente a massa tumoral, contudo eliminaria o foco de renovação do tumor e diminuiria lentamente o seu volume, reduzindo drasticamente sua capacidade de recidiva (Dalerba et al., 2007). À luz destes novos conhecimentos, os futuros tratamentos deveriam focar nessa subpopulação de células tumorais,

evitando o comprometimento das CTs saudáveis do indivíduo, visto que são especialmente importantes para a renovação dos tecidos lesados pelo câncer (KLONISCH et al., 2008).

O desenvolvimento desses novos tratamentos seria dependente da identificação de marcadores específicos para as CTs dos tecidos tumorais. O fator de transcrição OCT4 é um alvo potencial, uma vez que é responsável por manter as células em um estado indiferenciado. Recentemente, Hammachi et al. (2012) demonstraram que a inibição desse fator de transcrição permite o desencadeamento do processo de diferenciação das células. O OCT4 possui um domínio de ligação ao DNA que possibilita o reconhecimento de uma sequência de nucleotídeos octamérica (ATGCAAAT) presente nas regiões regulatórias de diversos genes relacionados à diferenciação celular. Um estudo, utilizando imunoprecipitação de cromatina (ChIP), identificou 623 possíveis associações do OCT4 com regiões promotoras em células tronco embrionárias (CTEs) (Wei et al., 2007). Tais resultados demonstram que o OCT4 atua na regulação de um grande número de genes, sendo um elemento central não só no desenvolvimento embrionário, mas, também, na manutenção dos estoques de CTs nos tecidos adultos. O OCT4, tanto inibe, quanto ativa genes necessários para pluripotência, podendo induzir ou inibir a diferenciação. O que define o seu modo de atuação pode ser a associação com outros fatores de transcrição tais como: SOX2, NANOG (WALKER et al., 2007) e FOXD3 (Liu e Labosky, 2008), ou modificações pós-traducionais da proteína (Plath, et al. 2009; Wang e Dai, 2010; Zhang et al., 2010)

O gene OCT4, também conhecido como POU5F1, pertencente à família POU e, em humanos, está localizado na posição p21.3 do cromossomo 6

(Wang e Dai, 2010). Este gene codifica para uma proteína que contém um domínio N-terminal (ND) do aminoácido 1 ao 133, um domínio POU do tipo homeobox (134-289) e um domínio C-terminal (CD), do aminoácido 290 ao 360 (LIM et al., 2009). Possui, o mesmo, pelo menos 3 isoformas de transcritos (OCT4A, OCT4B e OCT4B1), sendo o OCT4A responsável pelas características-tronco. A função do OCT4B1 ainda é desconhecida, apesar de se acreditar que também atue como fator tronco (Wang e Dai, 2010). A isoforma OCT4B contém entradas alternativas do ribossomo (IRES) que possibilitam a tradução para três isoformas proteicas OCT4B-265, OCT4B-190 e OCT4B-164, sendo a proteína OCT4B-190 relacionada ao stress celular (GAO et al., 2010). Além das isoformas citadas acima, o OCT4 apresenta também pseudogenes, os quais são caracterizados por sequências de DNA genômico similares aos genes e são considerados parentes extintos dos genes funcionais. Contudo, já foram descritos pseudogenes relacionados ao OCT4 que são transcritos *in vivo* e, potencialmente, podem atuar na regulação da isoforma principal do próprio OCT4 (Liedtke et al., 2007). Além da sua íntima relação com as CTs, o OCT4 também tem sido associado ao desenvolvimento de tumores. Fang et al., (2011) observaram uma superexpressão deste gene em diversos tipos de câncer e relacionaram essa característica com o conceito de célula tronco tumoral. Dessa forma, o OCT4 pode ser considerado como um marcador específico de CTTs, conforme o demonstrado por Looijenga et al., (2003) através de marcação imunohistoquímica em diversos tipos de células tumorais germinativas.

Entre os inúmeros genes controlados pelo OCT4, encontram-se aqueles que codificam para as proteínas da superfamília ABC (*ATP binding cassette*).

Essa superfamília é composta por aproximadamente 300 proteínas transmembrana que atuam transportando diferentes compostos para fora das células de uma forma ATP-dependente (Stavrovskaya e Stromskaya, 2008). Essas proteínas, quando em alta concentração na membrana, são responsáveis pela maior falha no tratamento de neoplasias, representando um fenótipo de resistência a múltiplas drogas conhecido como MDR. Foram encontrados diversos tipos cancerígenos resistentes a quimioterápicos, como o de células escamosas orais (Pérez-Sayáns et al., 2010; Zhu et al., 2012), carcinoma gástrico (BORSKA et al., 2012), próstata (Bhangal et al., 2000; Kato et al., 2012) mama e pulmão (Amiri-Kordestani et al., 2012) pâncreas (Lee, et al., 1996) dentre outros.

Alguns fenômenos associados a transcrição dos genes de resistência, podem comprometer ou elevar os níveis transcricionais, um deles é a metilação, a qual pode silenciar um gene através de proteínas de ligação metil-CpG que atuam em um processo de unificação do silenciamento da transcrição por associação bioquímica com desacetilases de histonas no caso dos genes MDR, a hipometilação pode tornar regiões do DNA mais expostas como o promotor das proteínas ABC, uma vez que células MDR tem o promotor do gene ABCB1 hipometilado enquanto que células não MDR tem o promotor hipermetilado (BAKER; EL-OSTA, 2003). Vários tipos cancerígenos apresentaram relação entre a hipometilação e resistência a múltiplas drogas como: câncer de próstata, mama, broncoalveolar, gástrico, carcinomas de cólon e leucemias (SHI et al., 2011).

(BERTOLINI et al., 2009) demonstraram que a expressão de OCT4 tem papel fundamental na resistência oferecida ao quimioterápico cisplatina, que atua como modulador das proteínas de resistência, em células de câncer pulmonar. Um estudo mais recente aponta para uma associação direta entre CTTs e o fenótipo MDR, mostrando uma relação entre a superexpressão de certos genes da superfamília ABC e o aumento na expressão do gene OCT4, de uma linhagem eritroleucêmica MDR resistente ao quimioterápico vincristina (VCR) (Marques et al., 2010) que atua como inibidor das proteínas de resistência. Entretanto, esses mesmos autores observaram que, tanto a linhagem MDR, quanto a parental não MDR, apresentaram o mesmo perfil de marcação CD34+/CD38-, evidenciando que ambas as linhagens apresentam características-tronco. Por essa razão, não foi possível estabelecer uma relação direta entre a expressão do OCT4 e o fenótipo MDR, a não ser que existam isoformas alternativas de OCT4 com distintas capacidades de regulação gênica, caracterizando subpopulações de CTTs com respostas diferenciadas frente à exposição ao quimioterápico.

Dentro do cenário descrito acima, o presente trabalho testou a hipótese de que podem existir isoformas específicas de OCT4 relacionadas ao fenótipo MDR. Para isso, foi utilizada a linhagem celular K562 de LMC. De origem hematopoiética, essa linhagem corresponde a uma Eritroleucemia humana que apresenta mieloblastos como tipo celular, com morfologia linfoblástica, correspondente a uma LMC em fase aguda ou blástica. Nessa fase ocorre intensa divisão celular, diferentemente da LMC em sua fase inicial que é lenta. estabelecida por Lozzio e Lozzio (1975). Uma segunda linhagem celular para estudo da LMC foi estabelecida por Rumjanek et al. (1994; 2001) utilizando

VCR para induzir uma linhagem eritroleucêmica resistente conhecida como K562-Lucena (Lucena), a qual já teve seu fenótipo MDR comprovado em vários estudos (Maia et al.,1996a; 1996b; Marques-Silva,1996; Orind et al.,1997). O objetivo desta pesquisa foi estudar fenômenos que levam a resistência à quimioterapia, comparando as linhagens tumorais leucêmicas K562 e sua parental MDR Lucena. Para isso, foi analisado o perfil de expressão de alguns genes dos transportadores ABC e de fatores de transcrição associados a eles. Também foi sequenciada a região proximal de genes ABC vinculados ao fenótipo MDR, bem como a própria região codificante do gene OCT4. Assim sendo, é possível que alterações no OCT4 ou nos seus sítios de reconhecimento nas regiões regulatórias dos genes ABC possam estar intimamente relacionadas à resposta apresentada pelas células MDR ao quimioterápico VCR previamente observada.

### **Objetivo geral**

O objetivo geral da presente Dissertação foi avaliar a relação entre o gene OCT4 e os fenômenos que levam à resistência a múltiplas drogas (MDR) nas linhagens eritroleucêmicas K562 e sua derivada resistente K562-Lucena1.

### **Objetivos específicos**

- 1) Avaliar a expressão dos genes que codificam para os principais transportadores relacionados ao fenótipo MDR e dos fatores de transcrição associados, na presença ou ausência do agente quimioterápico.
- 2) Verificar se existem mutações nas regiões proximais dos promotores dos principais transportadores ABC, comparando as linhagens K562 e Lucena.



- 3) Mapear os potenciais sítios de ligação do OCT4 nos promotores dos genes que codificam para os fatores de transcrição OCT4, SOX2, NANOG e FOXD3.
- 4) Isolar e sequenciar o OCT4 das linhagens K562 e Lucena, comparando com as sequências de OCT4 humano disponíveis no GenBank.
- 5) Verificar se as mutações encontradas implicam substituições de aminoácidos.
- 6) Analisar se as substituições de aminoácidos representam mudanças no padrão de fosforilação da proteína.

## **Material e Métodos**

### *Células e condições de cultura*

As células K562 (sensíveis) e Lucena (MDR) foram obtidas do Laboratório de Imunologia Tumoral, no Departamento de Bioquímica Médica da Universidade Federal do Rio de Janeiro (UFRJ). As duas linhagens foram cultivadas em meio RPMI-1640 (GIBCO), suplementado com 10% de soro fetal bovino (GIBCO), bicarbonato de sódio (0,2 g/L), glutamina (0,3 g/L), HEPES (3 g/L), 1% de antibiótico e antimicótico, em garrafas de cultura, à 37°C. Além disso, a linhagem resistente recebeu VCR (Sigma, Brasil) a 60nM para manter o fenótipo MDR.

Para avaliar a viabilidade celular (percentual de células viáveis pelo número de células totais) foi utilizado o método de exclusão por azul de Trypan. Esse método consiste em diluir a suspensão celular ( $2 \times 10^6$  células/ml) com uma solução de azul de trypan (0,08%), na proporção 1:1 e colocar essa

solução em câmara de Neubauer para contagem (células azuis representam células não íntegras, células não coradas representam células viáveis).

#### *Extração do RNA total e síntese de cDNA*

O RNA Total foi extraído de cinco amostras coletadas de cada linhagem celular com a quantidade de  $3 \times 10^6$  células por amostra de células de acordo com o protocolo do fabricante do reagente Trizol (Invitrogen, Brasil) Foram utilizadas 3 garrafas de cultivo com 10 ml de meio RPMI contendo cada uma aproximadamente  $300 \times 10^4$  células, sendo estas colocadas em um mesmo tubo falcon após centrifugação e assim realizado o cálculo para obter a concentração de  $3 \times 10^6$  células por amostra. Os RNAs foram quantificados por fluorimetria em um fluorímetro Qubit (Invitrogen, Brasil), utilizando o kit de análise Quanti-iT RNA (Invitrogen, Brasil). A integridade do RNA foi confirmada por eletroforese em gel de agarose a 1% com 0,5 µg/mL de brometo de etídio. Os RNAs foram tratados com Dnase 1 (Invitrogen, Brazil) e transcritos reversamente para cDNA (DNA complementar), utilizando 2 µg do RNA total, como descrito no protocolo do fabricante da enzima transcriptase reversa (RT High capacity, Invitrogen, Brasil).

#### *PCR quantitativo em tempo real*

A análise da expressão gênica foi realizada utilizando o equipamento Real-time PCR System 7500 (Applied Biosystems, Brasil). Cada reação de PCR teve 10 µL de volume final, contendo 5 µL de SYBR Green PCR Master Mix (Applied Biosystems, Brasil), 0,2 µL de ROX 0,25 µL de cada *primer*, 2,7 µL de água ultra pura (Invitrogen, Brasil) e 1,6 µL de cDNA concentrado das linhagens

celulares K562 ou Lucena. Os *primers* utilizados neste estágio foram desenhados com auxílio da ferramenta Primer-Blast do website NCBI (<http://www.ncbi.nlm.nih.gov>). Os genes analisados foram da superfamília ABC (ATP Binding Cassette) e dos fatores de transcrição OCT4, NANOG, SOX2 e FOXD3, a partir de sequências disponíveis no GenBank. O gene EF1 $\alpha$  foi utilizado como controle endógeno para normalização de dados. Informações referentes às sequências dos *primers* utilizados para executar a expressão gênica estão presentes na Tabela 1. Os dados de expressão gênica foram analisados comparando a mediana $\pm$ DP entre as linhagens celulares através do programa REST (*Relative Expression Software Tool*), de acordo com Pfaffl *et al.* (2002). REST utiliza o método  $2^{-\Delta\Delta Ct}$  (Livak & Schmittgen, 2001) seguido de um método de randomização (Pair Wise Fixed Reallocation Randomization Test) para testar a diferença significativa dos índices calculados de expressão e para gerar um erro padrão estimado é utilizado o algoritmo de Taylor.

**Tabela 1.** *Primers (forward e reverse)* utilizados na expressão gênica, com seus respectivos acessos no GenBank.

| Gene         | Primers (5'-3')  | GenBank (número de acesso) |
|--------------|--|----------------------------|
| EF1 $\alpha$ | Forward: GCCAGTGGAAACCACGCTGCT<br>Reverse: ATCCTGGAGAGGCAGGGCGCA         | CCDS4980                   |
| ABCB1        | Forward: TCCTCAGTCAAGTTCAGAGTCTTCA<br>Reverse: TCTCCACTTGATGATGTCTCTCACT | CCDS5608                   |
| ABCG2        | Forward: GCAGGGACGAACAATCATCT<br>Reverse: TTCCTGAGGCAATAAGGTG            | CCDS3628                   |
| ABCC1        | Forward: CAGACTGGCAGGGCTACTTC<br>Reverse: CTCACGAAGCAGATGTGGAA           | CCDS42122                  |
| ABCC2        | Forward: AGAGCTGGCCCTTGTACTCA<br>Reverse: TGCCTAGGTAGAGGCTTGGA           | CCDS7484                   |
| ABCC3        | Forward: TGTGCTAGCTGATGGACAGG<br>Reverse: GCAGAGAAAGTTGGCAAAGG           | CCDS32681                  |

|        |  |           |
|--------|--|-----------|
| ABCC4  | Forward: AGGCTTCTGTGCGTCATTCT<br>Reverse: CTTGGAGAGGAGTTGCAAGG   | CCDS9474  |
| ABCC11 | Forward: ACCGCATCCTGGTTATGGGCA<br>Reverse: GGCCATGAGGGCTGCGAACAA | CCDS10732 |
| OCT4   | Forward: TTCCCATGGCGGGACACCT<br>Reverse: CCCCTGGCCCATCACCTCCA    | CCDS34391 |
| NANOG  | Forward: GGTGTGACGCAGAAGGCCTCA<br>Reverse: AGTCGGGTTACACAGGCATCC | CCDS31736 |
| SOX2   | Forward: AAAAACAGCCCGGACCGCGT<br>Reverse: CTCCTGGGCCATCTTGCGCC   | CCDS3239  |
| FOXD3  | Forward: CTCCCCGGATCCCGACTGG<br>Reverse: GGACGGCCACTTCGGGGTTTT   | CCDS624   |

### *Tratamento com vincristina (VCR)*

Para realizar o primeiro experimento de expressão gênica, a linhagem resistente a múltiplas drogas Lucena permaneceu sem receber VCR por 45 dias e, para o segundo, foi exposta a ele por quatro meses. Logo após a cada um dos tratamentos (com e sem VCR) foi extraído o RNA dessas células para análises de expressão gênica.

### *Sequenciamento dos promotores gênicos ABCB1 e ABCG2*

O DNA genômico das duas linhagens celulares foi extraído através do kit Axyprep (Axygen Biosciences, Brasil), utilizando enzima RNase (Invitrogen, Brasil) para purificação do DNA. A qualidade do DNA extraído foi verificada por eletroforese em gel de agarose 1%. Os genes ABCB1 e ABCG2 tiveram seus promotores sequenciados. Dois pares de *primers* específicos para amplificar, no total, cerca de 1.400 pb de cada promotor foram desenhados manualmente a partir das sequências disponíveis no GenBank. As reações de PCR foram conduzidas com as seguintes condições: desnaturação inicial de 2' a 94°C; 35 ciclos de 30" a 94°C; 30" a 60°C e 2' a 72°C e uma extensão final de 10' a

72°C. Os produtos obtidos foram submetidos à eletroforese em gel de agarose 1% para verificar a qualidade e concentração do produto amplificado. O Kit Illustra GFX (GE Healthcare, Brasil) foi utilizado para purificação dos produtos de PCR, os quais foram sequenciados diretamente em sequenciador automático MegaBACE 500 (GE Healthcare, Brasil). Os *primers* utilizados para ambos os promotores e os acessos no GenBank das sequências utilizadas estão descritos na Tabela 2.

**Tabela 2.** *Primers (forward e reverse)* utilizados para o isolamento e sequenciamento dos promotores ABCB1 e ABCG2, com seus respectivos acessos no GenBank.

| Gene  | <i>Primers (5'-3')</i>  | GenBank<br>(número de acesso) |
|-------|---|-------------------------------|
| ABCB1 | ABCB1(P1)-FOR: AGCATGCTGAAGAAAGACCACTGC<br>ABCB1(P1)-REV: CCCCACTTGGTCCCATGGACT   | NG_011513                     |
|       | ABCB1(P2)-FOR: CCGTTCCAAGGAGCGCGAGG<br>ABCB1(P2)-REV: TGCTGTGCTCCACTCAGCCA        |                               |
| ABCG2 | ABCG2(P1)-FOR: AGGAGGACTGTTTGGGAACAGCA<br>ABCG2(P1)-REV: TGCAGGGACAGACTACATGCTCA  | NG_032067                     |
|       | ABCG2(P2)-FOR: TCTCCTGTTCTGACTGCTTGCC<br>ABCG2(P2)-REV: GCTCCTTCAGTAAATGCCTTCAGGT |                               |

#### *Alinhamento dos promotores sequenciados*

As sequências de nucleotídeos dos promotores gênicos das duas linhagens celulares foram submetidas à comparação com as sequências disponíveis no GenBank, usando a ferramenta de busca BLAST (*Basic Local Alignment Search Tool*). Essa ferramenta permitiu a confirmação de que o promotor isolado realmente é aquele correspondente ao presente no genoma humano. Adicionalmente, foi realizada uma comparação de cada promotor entre a sequências do GenBank e aquelas isoladas das linhagens K562 e

Lucena, através de alinhamento. O percentual de eventuais mutações encontrado nos promotores sequenciados foi verificado sempre em relação às sequências disponíveis no GenBank.

#### *Mapeamento dos sítios de ligação do OCT4 nos promotores gênicos*

Os promotores foram analisados através do programa MatInspector (Quandt *et al.*, 1995). Essa ferramenta utiliza uma grande base de dados de sítios de ligação ao DNA reconhecidos por fatores de transcrição. A cada sítio encontrado para um determinado fator de transcrição, um nível de similaridade é atribuído facilitando a filtragem dos dados e a seleção de sítios com maior potencial de ligação. Os sítios de ligação ao OCT4 foram identificados não somente nos promotores sequenciados (ABCB1 e ABCG2), mas, também, nos promotores do OCT4 (NC\_000006.11), SOX2 (NC\_000003.11), NANOG (NC\_000012.11) e FOXD3 (NC\_000001.10) disponíveis no GenBank.

#### *Sequenciamento da região codificante do gene OCT4 de K562 e Lucena*

Para isolamento da região codificante do gene OCT4 foram realizadas PCRs, utilizando cDNA de cada linhagem celular como molde, com as seguintes condições: desnaturação inicial de 2' a 94°C; 35 ciclos de 30" a 94°C; 30" a 60°C e 2' a 72°C e uma extensão final de 10' a 72°C. *Primers* específicos foram desenhados manualmente nas regiões 5' e 3' UTR do OCT4 (hOCT4-5UTR: 5'-CTTGGGGCGCCTTCCTTCCC-3'; hOCT4-3UTR: 5'-AGCTCCTCCCCTCCCCCTGT-3'), com o objetivo de amplificar a região codificante completa com aproximadamente 1.150 nucleotídeos. Os produtos obtidos foram submetidos à eletroforese em gel de agarose 1% para verificar a

qualidade e concentração do produto amplificado. O Kit Ilustra GFX (GE Healthcare, Brasil) foi utilizado para purificação dos produtos de PCR, os quais foram diretamente sequenciados em sequenciador automático MegaBACE 500 (GE Healthcare, Brasil).

As sequências de nucleotídeos dos gene OCT4 das duas linhagens celulares foram submetidas à comparação com as sequências disponíveis no GenBank, usando a ferramenta de busca BLAST (*Basic Local Alignment Search Tool*). Essa ferramenta permitiu a confirmação de que o gene isolado realmente é aquele correspondente ao presente no transcriptoma humano. Posteriormente, foi realizada uma comparação desse gene entre a sequência do GenBank e aquelas isoladas das linhagens K562 e Lucena, através do alinhamento das sequências. O percentual de eventuais mutações encontradas no OCT4 foi verificado sempre em relação às sequências disponíveis no GenBank. Para ilustrar as diferenças entre as sequências obtidas, foi elaborada uma árvore de distância com os dados desse alinhamento, utilizando o índice de distância de Tamura-Nei (Tamura e Nei, 1993) e o método de Neighbor-joining para construção de árvore.

#### *Alinhamento das sequências deduzidas de aminoácidos de OCT4 e análise do padrão de fosforilação*

Para identificar as possíveis substituições de aminoácidos foi realizado um alinhamento da sequência de aminoácidos deduzidos do OCT4 de ambas as linhagens K562 e Lucena, em comparação com o OCT4 humano disponível no GenBank. Além disso, analisamos os potenciais motivos de fosforilação, utilizando a ferramenta de busca por motivos de fosforilação (*Phospho-motif*

*Finder*) encontrada no site do banco de dados *Human Protein Reference Database* (<http://www.hprd.org/>) (Peri et al., 2003).

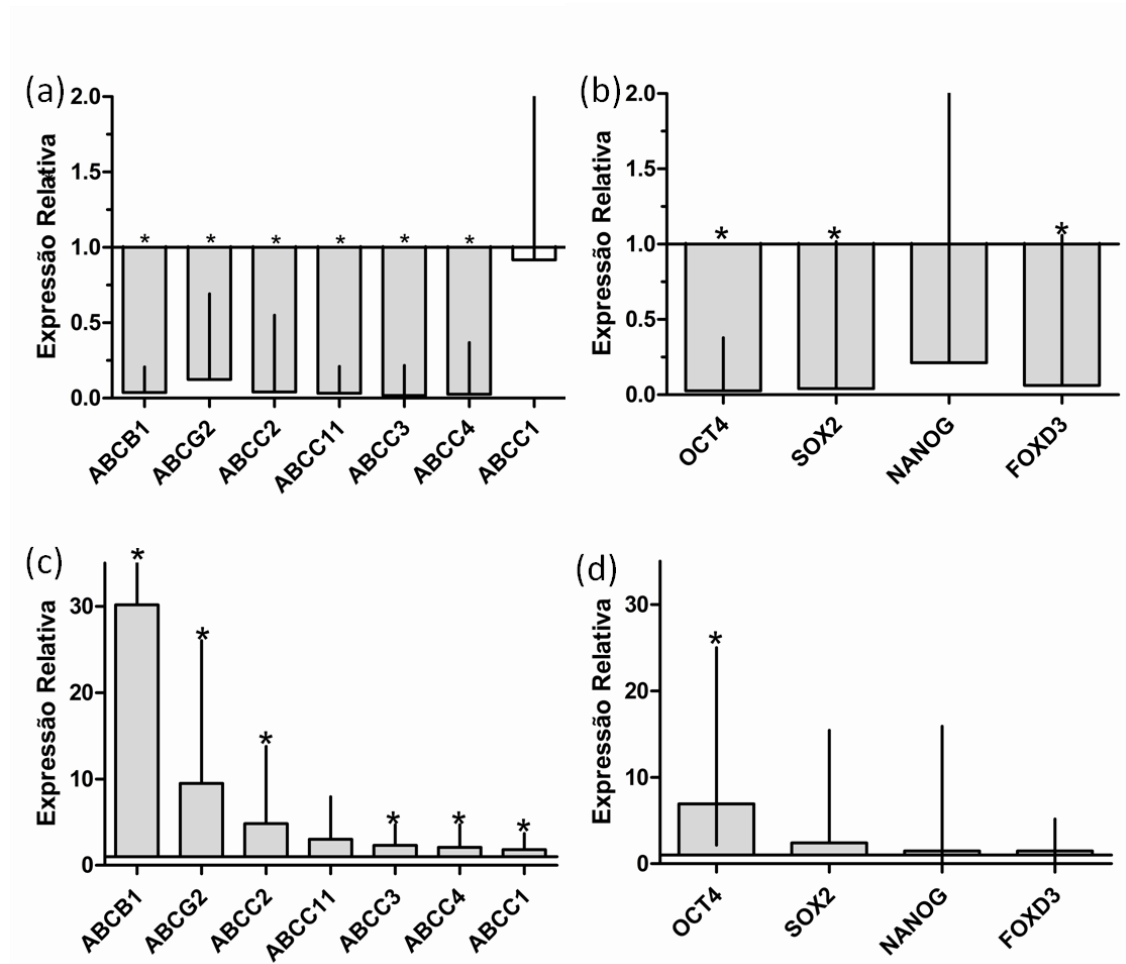
## **Resultados**

### *Análises de expressão gênica*

Conforme as Figuras 1a e 1b, as células da linhagem Lucena que não receberam VCR apresentaram uma diminuição significativa na expressão, tanto nos genes da superfamília ABC, quanto nos fatores de transcrição. Apenas ABCC11 e NANOG não tiveram diferença em relação à linhagem não MDR K562.

Após a linhagem Lucena receber VCR por quatro meses, foi observado que todos os genes ABC apresentaram um aumento significativo na expressão. O gene ABCB1 teve uma indução de 30 vezes, o ABCG2 de 9.5 vezes e o ABCC2 de 5 vezes. Os demais também tiveram expressão aumentada com valores menores. Apenas ABCC11 não mostrou diferença significativa (Figura 1c). Quanto aos fatores de transcrição, SOX2, NANOG e FOXD3 não apresentaram diferenças significativas. Contudo, OCT4 apresentou uma expressão elevada na linhagem Lucena da ordem de 7 vezes em relação à linhagem K562 (Figura 1d).



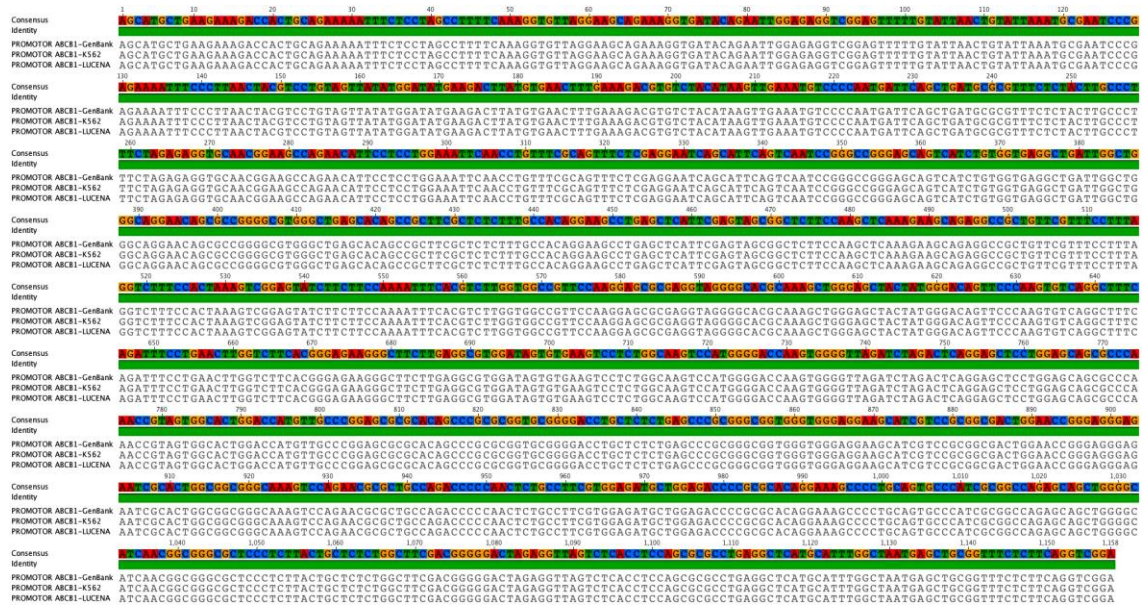


**Figura 1.** Expressão gênica de sete transportadores da superfamília ABC e dos fatores de transcrição OCT4, SOX2, NANOG e FOXD3 em células da linhagem Lucena (MDR) em relação à sua parental não MDR (K562). (a) e (b) Lucena sem o quimioterápico vincristina por 45 dias (VCR); (c) e (d) Lucena tratada com VCR por quatro meses. A expressão de todos os genes na linhagem K562 foi considerada = 1, representada pela linha horizontal paralela ao eixo X. Asteriscos significam diferença significativa estatisticamente ( $p < 0.05$ ).

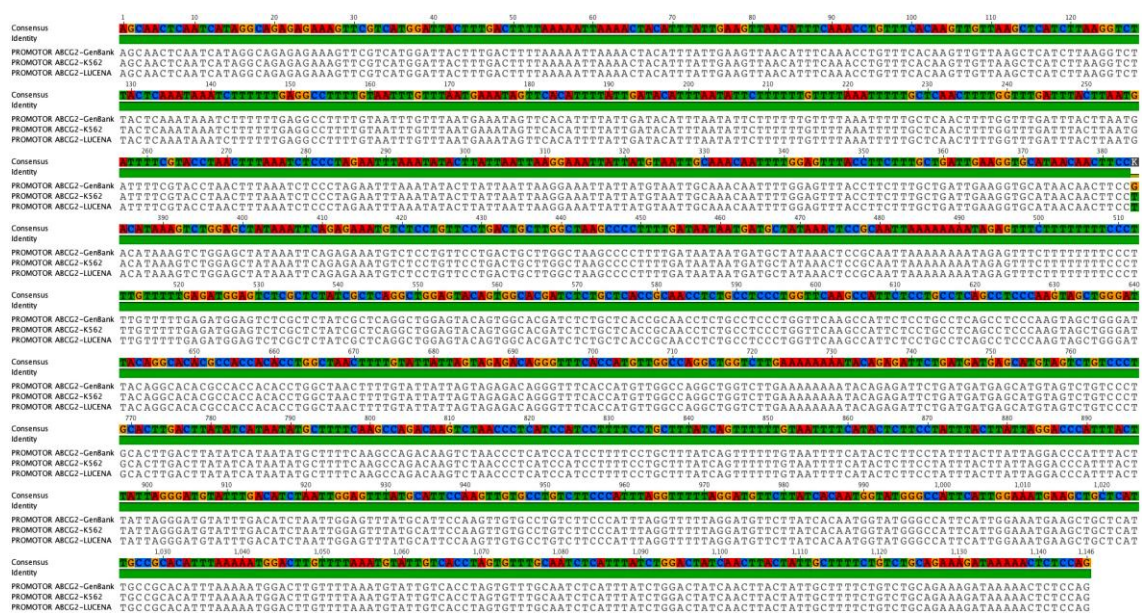
### Sequenciamento dos promotores ABCB1 e ABCG2

As regiões proximais dos promotores dos genes ABCB1 e ABCG2 das linhagens K562 e Lucena foram alinhadas com as regiões correspondentes

provenientes do GenBank. Os alinhamentos resultantes estão representados na Figura 2 (ABCB1) e na Figura 3 (ABCG2), onde apenas uma mutação foi encontrada em ambas as linhagens celulares na posição -384 do promotor do gene ABCG2.



**Figura 2.** Alinhamento de 1.158 nucleotídeos da região proximal do gene ABCB1, comparando as linhagens K562 e Lucena com a sequência correspondente do genoma humano disponível no GenBank.



**Figura 3.** Alinhamento de 1.146 nucleotídeos da região proximal do gene ABCG2, comparando as linhagens K562 e Lucena com a sequência correspondente do genoma humano disponível no GenBank.

*Mapeamento dos sítios de ligação do OCT4 nos promotores gênicos*

Através do programa MatInspector, foi possível analisar os potenciais sítios de ligação do OCT4 nos promotores dos fatores de transcrição (OCT4, SOX2, NANOG e FOXD3). A Tabela 3 mostra os resultados dessa análise, destacando o grande número de potenciais sítios de ligação do OCT4 em seu próprio promotor (nove sítios) e SOX2 (13 sítios), enquanto somente sete foram detectados em NANOG e quatro em FOXD3. A similaridade dos potenciais sítios de ligação do OCT4, em relação ao octâmero ATGCAAAT, variou de 0,783 a 0,998. A média de similaridade, para todos os sítios de SOX2, foi de 0,907, enquanto que para OCT4 foi 0,836, para NANOG 0,886 e para FOXD3 0,875.

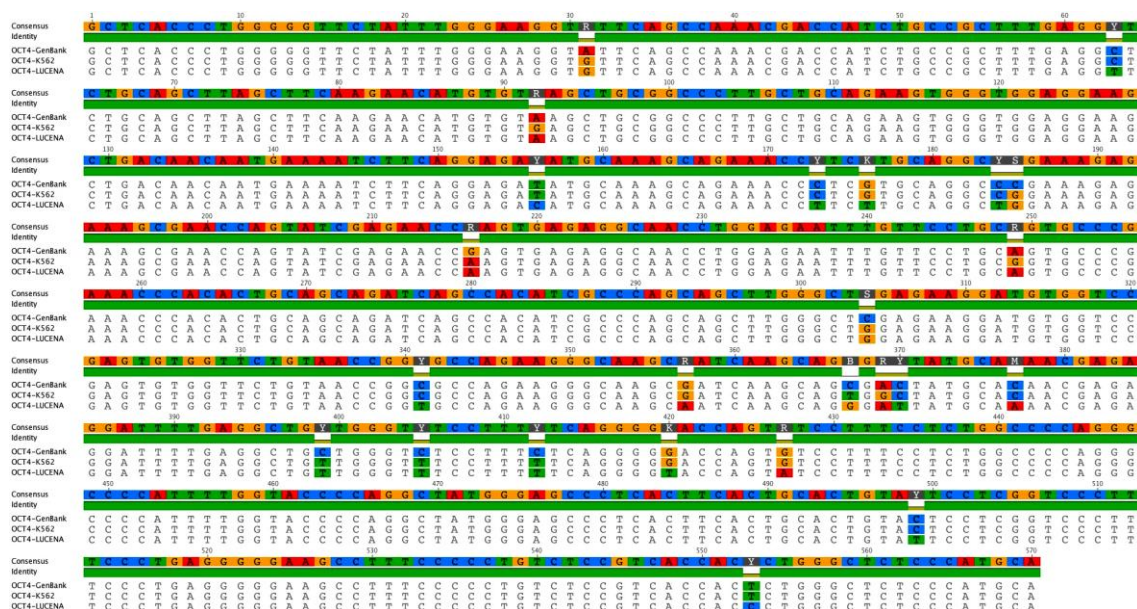
**Tabela 3.** Mapeamento dos sítios de ligação do fator de transcrição OCT4 nos promotores SOX2, OCT4, NANOG e FOXD3, evidenciado a posição de ligação, a fita de DNA e a similaridade à sequência octamérica ATGCAAAT.

| SOX2    |      |              | OCT4    |      |              | NANOG   |      |              | FOXD3   |      |              |
|---------|------|--------------|---------|------|--------------|---------|------|--------------|---------|------|--------------|
| Posição | Fita | Similaridade | Posição | Fita | Similaridade | Posição | Fita | Similaridade | Posição | Fita | Similaridade |
| 795     | +    | 0,887        | 24      | +    | 0,783        | 1133    | +    | 0,833        | 296     | -    | 0,984        |
| 831     | -    | 0,945        | 555     | -    | 0,845        | 1321    | -    | 0,816        | 941     | +    | 0,805        |
| 837     | -    | 0,951        | 743     | -    | 0,951        | 1419    | +    | 0,829        | 1509    | -    | 0,828        |
| 865     | +    | 0,806        | 1089    | +    | 0,809        | 2580    | +    | 0,857        | 3664    | -    | 0,884        |
| 1140    | -    | 0,812        | 1090    | -    | 0,79         | 3154    | -    | 0,927        |         |      |              |
| 1482    | -    | 0,975        | 2506    | +    | 0,87         | 3501    | +    | 0,941        |         |      |              |
| 1564    | -    | 0,985        | 2709    | -    | 0,807        | 4673    | +    | 0,998        |         |      |              |
| 1567    | +    | 0,906        | 3124    | -    | 0,853        |         |      |              |         |      |              |
| 1759    | -    | 0,901        | 3155    | +    | 0,813        |         |      |              |         |      |              |
| 2000    | -    | 0,914        |         |      |              |         |      |              |         |      |              |

|      |   |       |
|------|---|-------|
| 2071 | - | 0,834 |
| 2527 | + | 0,941 |
| 4466 | - | 0,935 |

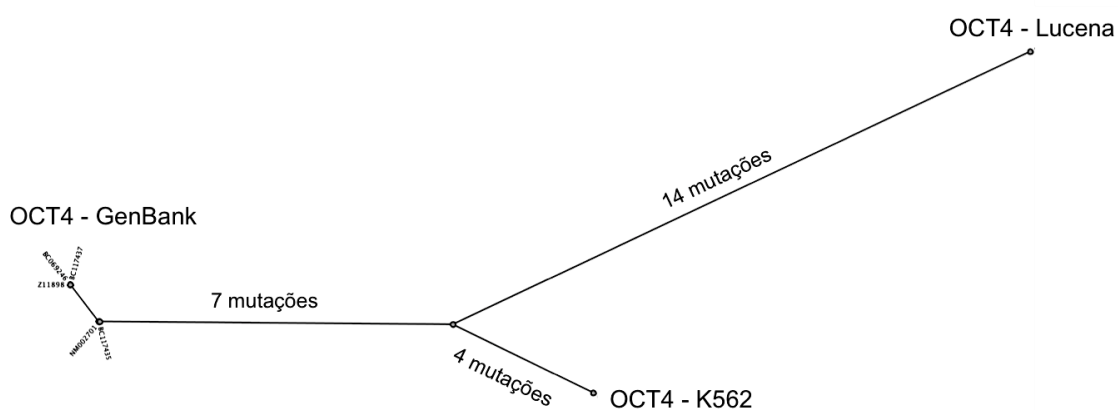
### Sequenciamento da região codificante do OCT4 em K562 e Lucena

Foram sequenciados 570 nucleotídeos da região contendo o domínio de ligação ao DNA do tipo homeobox, presente no gene OCT4 nas linhagens K562 e Lucena. As sequências foram analisadas e o alinhamento mostrado na Figura 4 revelou inúmeras mutações entre as linhagens. Em relação às sequências do GenBank, K562 apresentou 11 diferenças nucleotídicas. Quanto à Lucena, foram observadas 21 mutações realizando a mesma comparação. Adicionalmente, foram observadas quatro mutações exclusivas na linhagem K562 e 14 em Lucena.



**Figura 4.** Alinhamento de 570 nucleotídeos da região codificante do gene OCT4, comparando as linhagens K562 e Lucena com a sequência correspondente do gene OCT4 humano disponível no GenBank.

A árvore de distância da Figura 5 revelou que as linhagens K562 e Lucena estão distantes das sequências presentes no Genbank. Embora elas apresentem similaridade até certo nível (sete mutações compartilhadas), ambas possuem mutações exclusivas que as separam. A linhagem Lucena é a que mais se diferencia das originais, estando muito mais distante que sua parental, devido a um grande número de mutações exclusivas.

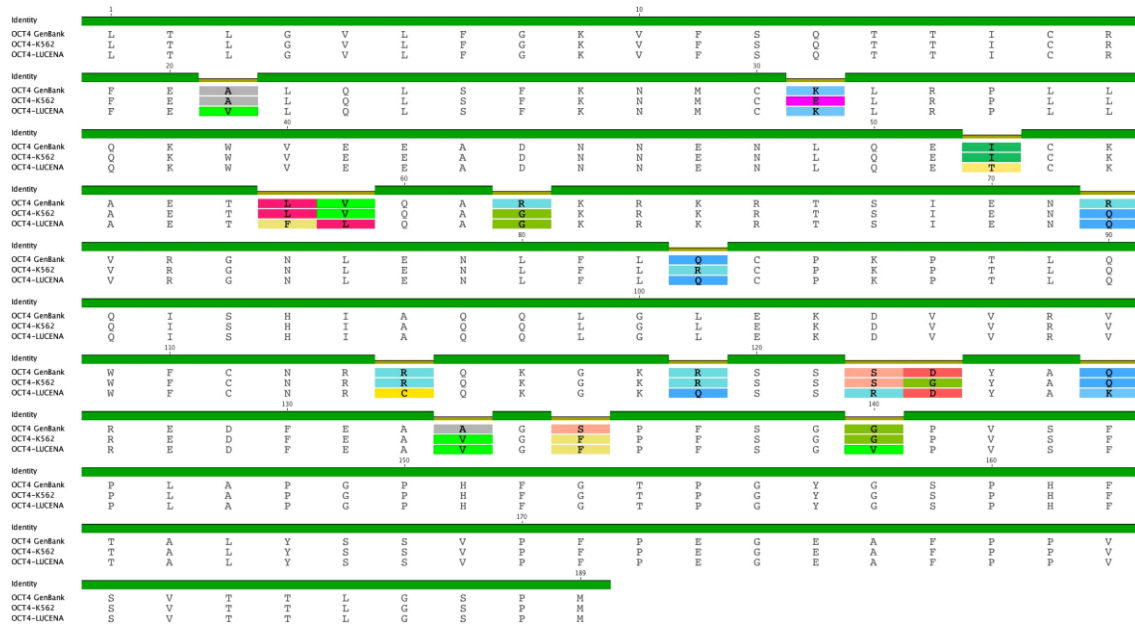


**Figura 5.** Árvore de distância entre os OCT4s humanos disponíveis no GenBank e os OCT4s das linhagens K562 e Lucena. O número de mutações compartilhadas e exclusivas está evidenciado nos ramos.

#### *Alinhamento dos aminoácidos do OCT4 e análise do padrão de fosforilação*

As sequências de aminoácidos deduzidas, tanto para a linhagem K562, quanto para Lucena, resultaram em 189 resíduos de aminoácidos. Essas sequências foram alinhadas com OCT4 presente no GenBank, com a intenção de encontrar possíveis substituições. Os resultados podem ser observados na Figura 6, revelando inúmeras substituições de aminoácidos em ambas as

linhagens. Quando comparadas ao OCT4 do GenBank, K562 apresentou sete substituições de aminoácidos (três exclusivas) e Lucena 13 (nove exclusivas).



**Figura 6.** Alinhamento de 189 aminoácidos deduzidos a partir das sequências nucleotídicas dos OCT4s das linhagens K562 e Lucena com a sequência correspondente do OCT4 humano disponível no GenBank.

Ao considerar esse alarmante número de substituições de aminoácidos, o passo seguinte foi realizar uma busca por potenciais motivos de fosforilação. Os resultados dessa análise *in silico* estão presentes na Figura 7. Pode ser observado que, quando comparado ao OCT4 do GenBank, K562 perdeu um motivo de fosforilação enquanto que Lucena perdeu dois deles.

| Protein Position | Sequence in the protein | Motifs (phosphorylated residues in red) | Phospho-motif types                      |
|------------------|-------------------------|---|--|
| 1 119 - 122      | RSSS                    | RXX <sup>o</sup> S                      | 14-3-3 domain binding motif              |
| 2 120 - 122      | SSS                     | S(pS/pT)X                               | MDC1 BRCT/ Prk1 PBD domain binding motif |
| 3 135 - 136      | SP                      | (pS/pT)P                                | WW domain binding motif                  |
| 4 154 - 155      | TP                      | (pS/pT)P                                | WW domain binding motif                  |
| 5 159 - 160      | SP                      | (pS/pT)P                                | WW domain binding motif                  |
| 6 167 - 169      | SSV                     | S(pS/pT)X                               | MDC1 BRCT/ Prk1 PBD domain binding motif |
| 7 187 - 188      | SP                      | (pS/pT)P                                | WW domain binding motif                  |



| Protein Position | Sequence in the protein | Motifs (phosphorylated residues in red) | Phospho-motif types                      |
|------------------|-------------------------|---|--|
| 1 119 - 122      | RSSS                    | RXX <sup>o</sup> S                      | 14-3-3 domain binding motif              |
| 2 120 - 122      | SSS                     | S(pS/pT)X                               | MDC1 BRCT/ Prk1 PBD domain binding motif |
| 3 154 - 155      | TP                      | (pS/pT)P                                | WW domain binding motif                  |
| 4 159 - 160      | SP                      | (pS/pT)P                                | WW domain binding motif                  |
| 5 167 - 169      | SSV                     | S(pS/pT)X                               | MDC1 BRCT/ Prk1 PBD domain binding motif |
| 6 187 - 188      | SP                      | (pS/pT)P                                | WW domain binding motif                  |



| Protein Position | Sequence in the protein | Motifs (phosphorylated residues in red) | Phospho-motif types                      |
|------------------|-------------------------|---|--|
| 1 120 - 122      | SSS                     | S(pS/pT)X                               | MDC1 BRCT/ Prk1 PBD domain binding motif |
| 2 154 - 155      | TP                      | (pS/pT)P                                | WW domain binding motif                  |
| 3 159 - 160      | SP                      | (pS/pT)P                                | WW domain binding motif                  |
| 4 167 - 169      | SSV                     | S(pS/pT)X                               | MDC1 BRCT/ Prk1 PBD domain binding motif |
| 5 187 - 188      | SP                      | (pS/pT)P                                | WW domain binding motif                  |



**Figura 7.** Motivos de fosforilação encontrados no gene OCT4 nas linhagens leucêmicas comparando com os motivos presentes no OCT4 humano disponível no GenBank. O domínio homeobox está delimitado pelo retângulo em preto e os grupos fosfato circulado representam as diferenças no padrão de fosforilação entre os OCT4s comparados. As tabelas associadas detalham a localização e os tipos de motivos de fosforilação detectados pela ferramenta *Phospho-motif Finder* (<http://www.hprd.org/>).

## Discussão

O fator de transcrição OCT4 é essencial no desenvolvimento embrionário e é requerido para a manutenção da totipotência e pluripotência das células tronco (CTs). Junto com outros fatores de transcrição como SOX2, NANOG e FOXD3, o OCT4 exerce sua função na manutenção da pluripotência através da ativação de genes que mantêm a célula indiferenciada e da repressão de outros que promovem a ontogenia (Guo et al., 2002; Kim et al., 2008). Recentemente, um estudo com uma linhagem eritroleucêmica resistente

ao quimioterápico vincristina (VCR) sugere uma relação entre o fenótipo de resistência a múltiplas drogas (MDR) e o OCT4 (Marques et al., 2010), possivelmente pela presença de células tronco tumorais (CTTs). No presente trabalho, foi testada a hipótese do relacionamento OCT4/MDR de uma forma mais abrangente, avaliando não somente a expressão de genes, mas, também, sequenciando a região do promotor proximal de dois importantes genes envolvidos na extrusão de quimioterápicos. Posteriormente, sequenciamos a região do domínio homeobox do gene OCT4 isolada das linhagens eritroleucêmicas K562 (não-MDR) e Lucena (MDR). Os resultados obtidos suportam uma relação direta entre a expressão do gene OCT4 e o fenótipo MDR, além de revelar, pela primeira vez, um grande número de mutações na região codificante do gene OCT4. Apesar dessas mutações aparecerem na linhagem não-MDR, elas ocorreram em dobro na linhagem resistente, sugerindo a presença de um OCT4 mutado, potencialmente relacionado ao fenótipo MDR.

O primeiro experimento realizado teve por objetivo comprovar o relacionamento entre a expressão do gene OCT4 e o fenótipo MDR em células eritroleucêmicas. Foram analisados o gene OCT4 e os fatores de transcrição associados SOX2, NANOG e FOXD3. Quanto ao fenótipo MDR, foram analisados os genes que codificam para os transportadores ABC mais relacionados à extrusão de quimioterápicos: ABCB1, ABCG2, ABCC1, ABCC2, ABCC3, ABCC4 e ABCC11 (Scotto, 2003). Os resultados mostrados na Figura 1 se referem à expressão dos fatores de transcrição e dos transportadores em duas situações: na presença ou na ausência de VCR. Foi observado que o aumento da expressão dos transportadores ocorre, concomitantemente, com o



aumento da expressão apenas de OCT4 nas células expostas a VCR. Esta observação comprova a relação desse fator de transcrição com o fenótipo MDR e levanta duas hipóteses: (1) existem mutações nos promotores dos genes ABC na linhagem MDR, permitindo um maior número de ligações do OCT4 e, conseqüentemente, aumento na transcrição; (2) existem mutações no próprio OCT4 da linhagem MDR, sendo que esta forma mutada encontraria mais sítios de ligação nos promotores ABC, aumentando sua transcrição.

Para testar a primeira hipótese, foram isolados e sequenciados aproximadamente 1.300 nucleotídeos do promotor proximal de dois dos mais importantes transportadores relacionados ao fenótipo MDR: ABCB1 e ABCG2. De fato, como mostrado na Figura 1b, esses dois genes foram os mais ativados com a presença da VCR na linhagem Lucena. A análise das sequências obtidas para os promotores ABCB1 e ABCG2, comparando as linhagens K562 e Lucena com a sequência equivalente disponível no GenBank, estão demonstradas nas Figuras 2 e 3, respectivamente. O resultado evidenciou apenas uma mutação no promotor ABCG2 em ambas as linhagens leucêmicas sendo, dessa forma, a hipótese 1 rejeitada.

Uma vez rejeitada a primeira hipótese, é possível que existam formas de OCT4 mutadas nas linhagens leucêmicas testadas. O resultado da análise de expressão dos fatores de transcrição, mostrado na Figura 1d, é uma evidência de que isso pode estar ocorrendo. Neste caso, a exposição à VCR induziu um aumento apenas no OCT4. Era esperado um aumento equivalente em SOX2, uma vez que a estequiometria destes dois fatores de transcrição é extremamente importante para que os genes relacionados sejam regulados de forma correta (Yu et al., 2007). Levando em consideração o estreito

relacionamento entre OCT4 e SOX2, é provável que ambos promotores sejam reconhecidos pelo OCT4. A análise *in silico* desses promotores e, também dos promotores dos demais fatores de transcrição relacionados, revelou o grau de interação entre eles e sua capacidade de autorregulação transcricional. Os resultados estão mostrados na Tabela 2, onde foram utilizados 5.000 nucleotídeos dos promotores OCT4, SOX2, NANOG e FOXD3 para a análise dos sítios de ligação de OCT4. Foram identificados nove sítios no promotor OCT4, 13 em SOX2, sete em NANOG e quatro em FOXD3. Isso evidencia a íntima relação, principalmente, entre OCT4 e SOX2, tendo o primeiro a capacidade de regular a sua própria transcrição. Porém, estes resultados não explicam a não indução da transcrição de SOX2 de forma equivalente a OCT4 na linhagem MDR exposta a VCR, ainda mais que a similaridade média dos 13 sítios para ligação de OCT4 no promotor SOX2 foi a maior observada (0,907). Uma justificativa plausível para essa observação seria a ocorrência de mutações no OCT4 da linhagem MDR, impedindo o reconhecimento dos sítios de ligação convencionais. Assim como é possível que mutações impeçam o reconhecimento dos sítios convencionais, existe a possibilidade de um OCT4 mutado reconhecer novos sítios de ligação em promotores que não são alvos usuais. Isso poderia explicar o aumento significativo observado na expressão dos transportadores ABCB1 e ABCG2 na linhagem MDR, visto que análises *in silico* destes promotores, realizadas por Marques et al. (2010) identificaram apenas dois sítios de ligação para OCT4 em cada um deles.

Para testar a segunda hipótese, um fragmento de 570 nucleotídeos contendo a região do domínio homeobox do gene OCT4 foi isolado e sequenciado para ambas as linhagem leucêmicas testadas. Essa região foi

escolhida por ser responsável pelo reconhecimento dos promotores-alvo através da sua capacidade de ancoragem aos sítios de ligação ao DNA. As sequências obtidas foram comparadas com a sequência correspondente disponível no GenBank, sendo que o alinhamento resultante está mostrado na Figura 4. Foram identificadas 24 mutações entre as três sequências comparadas. A linhagem K562, quando comparada apenas com a sequência do GenBank, mostrou 11 mutações, sendo quatro exclusivas da linhagem. A linhagem Lucena, quando comparada apenas com a sequência do GenBank, mostrou 21 mutações, sendo 14 exclusivas da linhagem. Esses resultados demonstram, pela primeira vez, que o gene OCT4 é um alvo para mutações, as quais podem estar relacionadas com o fenótipo MDR. Uma análise de distância entre as sequências aqui obtidas e sequências humanas de OCT4 normal disponíveis no GenBank, é mostrada na Figura 5. A árvore resultante demonstra que K562 está bem mais próxima das sequências normais de OCT4 do que Lucena. A questão que surge é: essas mutações implicam substituição de aminoácidos?

Mutações podem, eventualmente, acarretar substituição de aminoácidos com possíveis alterações de função na proteína resultante. Recentemente, Brumbaugh et al. (2012) realizaram uma análise funcional de dois aminoácidos dentro da região homeobox do OCT4 (T234 e S235) e observaram que a perda de sítio de fosforilação, induzida por mutação, implicou alteração na atividade regulatória deste fator de transcrição. Para avaliar se as mutações aqui observadas implicam substituição de aminoácidos, as sequências deduzidas de aminoácidos foram alinhadas com a proteína OCT4 disponível no GenBank. O alinhamento está mostrado Figura 6 e evidencia que, quando comparada ao

GenBank, K562 teve sete aminoácidos substituídos (três exclusivos). Por sua vez, Lucena teve 13 substituições, sendo nove exclusivas. Esses resultados comprovam que o OCT4 de K562 é diferente do OCT4 de Lucena e ambos são bastante diferentes do OCT4 disponível no GenBank. Sem dúvida, as mutações observadas implicaram substituição de aminoácidos em ambas as linhagens analisadas. Portanto, a nova questão é: essas substituições implicam alteração de função?

Em geral, substituições de aminoácidos podem afetar o padrão de processamento pós-translacional das proteínas. Com relação ao OCT4, alguns estudos especulam que o seu padrão de fosforilação controla sua atividade (Brehm et al., 1997; Saxe et al., 2009). A fosforilação é um dos principais mecanismos de regulação de atividade das proteínas, em que grupos fosfato são transferidos por enzimas quinases para aminoácidos específicos em eucariotos: serina, treonina e tirosina. Para avaliar se as substituições de aminoácidos observadas no OCT4 de K562 e Lucena resultam em mudanças no padrão de fosforilação deste fator de transcrição, foi realizada uma análise de identificação dos sítios potenciais de fosforilação. O resultados dessa análise estão demonstrados na Figura 7, em que se pode observar que K562 perdeu um sítio de fosforilação, enquanto Lucena perdeu dois, quando comparadas com o OCT4 do GenBank. Esta observação pode ter uma implicação profunda na atividade do OCT4 de K562 e, talvez mais especialmente, no OCT4 de Lucena. A forma com que essas diferenças podem atuar na regulação da atividade de cada OCT4 não podem ser inferidas, porém é possível que estejam relacionadas com o conceito de célula tronco tumoral e com o fenótipo MDR.

Os resultados obtidos neste estudo demonstram, que o gene OCT4 está relacionado com o fenótipo MDR e é um alvo de mutações nas linhagens eritroleucêmicas analisadas. Esse fato fica mais evidente, quando não encontramos mutações exclusivas nos promotores dos transportadores sequenciados. A taxa de mutação calculada para o OCT4 de K562 foi de uma a cada 52 bases, enquanto que para Lucena foi de uma a cada 27. Essas mutações implicaram substituições de aminoácidos com conseqüente alteração no padrão de fosforilação do domínio de ligação ao DNA deste fator de transcrição. Moléculas de OCT4, com diferentes padrões de fosforilação, podem controlar diferentes conjuntos de genes ou modificar sua relação com os genes naturalmente controlados. Estas informações podem abrir um novo campo para o tratamento do câncer, especialmente aqueles resistentes aos quimioterápicos, uma vez que a caracterização de formas diferenciadas de OCT4 pode permitir a identificação de alvos mais específicos para o desenvolvimento de novas drogas antineoplásicas para o combate a tumores resistentes à quimioterapia.

### **Conclusões**

- 1) O fator de transcrição OCT4 está relacionado com o fenótipo MDR na linhagem eritroleucêmica resistente K562-Lucena1 através da modulação da transcrição dos principais genes da superfamília de transportadores ABC.
- 2) Não existem diferenças entre os promotores dos principais genes ABC relacionados ao fenótipo MDR, comparando as linhagens K562 e Lucena, que justifiquem as alterações observadas nos perfis de expressão gênica.

- 3) O gene OCT4 é um alvo de mutações nas linhagens eritroleucêmicas K562 e Lucena.
- 4) A taxa de mutações no OCT4 é muito maior na linhagem MDR (Lucena) do que na não MDR (K562).
- 5) As mutações observadas implicaram substituições de aminoácidos com consequente alteração no padrão de fosforilação das proteínas sendo que, quando comparadas com o OCT4 normal, K562 perdeu um sítio de fosforilação enquanto que Lucena perdeu dois deles.
- 6) As alterações observadas no gene OCT4 podem abrir um novo campo para o tratamento não apenas da LMC, mas também para as formas resistentes aos quimioterápicos deste tipo de câncer.

### **Perspectivas**

- Sequenciar a região codificante completa do gene OCT4 e avaliar a existência de mais subtipos específicos deste fator de transcrição para as linhagens K562 e Lucena.
- Silenciar o gene OCT4 mutado de Lucena e avaliar se ocorre a perda do fenótipo MDR.
- Verificar se os OCT4 mutados poderiam controlar diferentes conjuntos de genes nas células eritroleucêmicas através do teste de imunoprecipitação de cromatina.
- Superexpressar o OCT4 característico de Lucena na linhagem K562 para verificar se haverá a aquisição do fenótipo MDR.

- Determinar se este fenômeno ocorre em outras linhagens tumorais resistentes de origens diferentes.
- Desenvolver anticorpos específicos para cada OCT4 identificado, com o objetivo de criar drogas antineoplásicas mais eficientes tanto para o tratamento da LMC quanto da LMC-MDR.

## Referências

- Alison, M. R., Lim, S. M. L., & Nicholson, L. J. (2011). Cancer stem cells: problems for therapy? *The Journal of pathology*, 223(2), 147-61. doi:10.1002/path.2793
- Bertolini, G., Roz, L., Perego, P., Tortoreto, M., Fontanella, E., Gatti, L., Pratesi, G., et al. (2009). Highly tumorigenic lung cancer CD133+ cells display stem-like features and are spared by cisplatin treatment. *Proceedings of the National Academy of Sciences of the United States of America*, 106(38), 16281-6. doi:10.1073/pnas.0905653106
- Bonnet D, Dick JE. (1997). Human acute myeloid leukemia is organized as a hierarchy that originates from a primitive hematopoietic cell. *Nature Medicine*, 3:730–37.
- Brehm A, Ohbo K, Schöler H (1997) The carboxy-terminal transactivation domain of Oct-4 acquires cell specificity through the POU domain. *Molecular Cell Biology* 17:154–162.
- Brumbaugh, J., Hou, Z., Russell, J. D., Howden, S. E., Yu, P., Ledvina, A. R., Coon, J. J., et al. (2012). Phosphorylation regulates human OCT4. *Proceedings of the National Academy of Sciences of the United States of America*, 109(19), 7162-8. doi:10.1073/pnas.1203874109.
- Dalerba, P., Cho, R. W., & Clarke, M. F. (2007). Cancer stem cells: models and concepts. *Annual review of medicine*, 58, 267-84. doi:10.1146/annurev.med.58.062105.204854

- Fang, X.-feng, Zhang, W.-yi, Zhao, N., Yu, W., Ding, D., Hong, X., Li, L.-sha, et al. (2011). Genome-wide analysis of OCT4 binding sites in glioblastoma cancer cells. *Journal of Zhejiang University. Science. B*, 12(10), 812-9. doi:10.1631/jzus.B1100059
- Gao, Y., Wang, X., Han, J., Xiao, Z., Chen, B., Su, G., & Dai, J. (2010). The novel OCT4 spliced variant OCT4B1 can generate three protein isoforms by alternative splicing into OCT4B. *Journal of genetics and genomics*, 37(7), 461-5.. doi:10.1016/S1673-8527(09)60065-5
- Guo Y, et al. (2002) The embryonic stem cell transcription factors Oct-4 and FoxD3 interact to regulate endodermal-specific promoter expression. *Proceedings of the National Academy of Sciences of the United States of America*, 99:3663–3667
- Hammachi, F., Morrison, G. M., Sharov, A. A., Livigni, A., Narayan, S., Papapetrou, E. P., O'Malley, J., et al. (2012). Transcriptional Activation by Oct4 Is Sufficient for the Maintenance and Induction of Pluripotency. *Cell Reports*, 1(2), 99-109. doi:10.1016/j.celrep.2011.12.002
- Hu, T., Liu, S., Breiter, D. R., Wang, F., Tang, Y., & Sun, S. (2008). Octamer 4 small interfering RNA results in cancer stem cell-like cell apoptosis. *Cancer research*, 68(16), 6533-40. doi:10.1158/0008-5472.CAN-07-6642
- Kim J, Chu J, Shen X, Wang J, Orkin SH (2008) An extended transcriptional network for pluripotency of embryonic stem cells. *Cell*, 132:1049–1061.
- Klonisch, T., Wiehac, E., Hombach-Klonisch, S., Ande, S. R., Wesselborg, S., Schulze-Osthoff, K., & Los, M. (2008). Cancer stem cell markers in common cancers - therapeutic implications. *Trends in molecular medicine*, 14(10), 450-60. doi:10.1016/j.molmed.2008.08.003
- Liedtke, S., Enczmann, J., Waclawczyk, S., Wernet, P., & Kögler, G. (2007). Oct4 and its pseudogenes confuse stem cell research. *Cell stem cell*, 1(4), 364-6. doi:10.1016/j.stem.2007.09.003



- Lim, H.-Y., Do, H.-J., Lee, W.-Y., Kim, D.-K., Seo, H. G., Chung, H.-J., Park, J.-K., et al. (2009). Implication of human OCT4 transactivation domains for self-regulatory transcription. *Biochemical and biophysical research communications*, 385(2), 148-53, doi:10.1016/j.bbrc.2009.05.029
- Livak, K. J., & Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2<sup>(-Delta Delta C(T))</sup> Method. *Methods (San Diego, Calif.)*, 25(4), 402–8. doi:10.1006/meth.2001.1262
- Liu, Y., & Labosky, P. a. (2008). Regulation of embryonic stem cell self-renewal and pluripotency by Foxd3. *Stem cells*, 26(10), 2475-84. doi:10.1634/stemcells.2008-0269
- Looijenga, L. H. J., Stoop, H., Leeuw, H. P. J. C. D., Tumors, C., Roozendaal, K. E. P. V., Zoelen, E. J. J. V., Weber, R. F. A., et al. (2003). POU5F1 ( OCT3 / 4 ) Identifies Cells with Pluripotent Potential in Human Germ Cell Tumors. *Cancer Research* , 63: 2244-2250.
- Lozzio, C.B.; Lozzio, B.B. (1975). Human chronic myelogenous leukemia cell-line with positive Philadelphia chromosome. *Blood*; 45(3): 321-334.
- Maia, R.C.; Silva, E.A.C.; Harab, R.C.; Lucena, M.; Pires, V.; Rumjanek, V.M..(1996a) Sensitivity of vincristine-sensitive K562 and vincristine-resistant K562-Lucena 1 cells to anthracyclines and reversal of multidrug resistance. *Brazilian Journal of Medical Biology Research*, 29(4): 467-472.
- Maia, R.C.; Wagner, K.; Cabral, R.H.; Rumjanek, V.M. (1996b). Heparin reverses Rhodamine 123 extrusion by multidrug resistant cells. *Cancer Letters*; 106: 101-108.
- Marques, D. S.; Sandrini, J. Z.; Boyle, R. T.; Marins, L. F.; Trindade, G. S. (2010). *Relationships between multidrug resistance (MDR) and stem cell markers in human chronic myeloid leukemia cell lines. Leukemia Research*; 34: 757–762.

- Marques-Silva, V.M. (1996). *Efeito de moduladores da diferenciação celular no processo de resistência a múltiplas drogas. Tese de Doutorado, Universidade Federal Fluminense, Niterói, RJ, Brasil.*
- Nguyen, L. V., Vanner, R., Dirks, P., & Eaves, C. J. (2012). Cancer stem cells: an evolving concept. *Nature reviews. Cancer*, 12(2), 133-43. doi:10.1038/nrc3184
- Orind, M.; Wagner-Souza, K.; Maia, R.C.; Rumjanek, V.M. (1997). *Modulation of P-glycoprotein on tumour cells. In: Sotelo JR & Benech JC (Ed.): calcium and cellular metabolism: transport and regulation. New York: Plenum Press, 72(1): 117-124.*
- Peri S, et al. (2003) Development of human protein reference database as an initial platform for approaching systems biology in humans. *Genome Research*, 13:2363–2371.
- Pfaffl, M.W.; Horgan, G.W.; Dempfle, L.,. (2002) Relative expression software tool (REST©) for group-wise comparison and statistical analysis of relative expression results in real-time PCR. *Nucleic Acids Research*, 30: 1–10.
- Plath, K., Huang, J., Saxe, J. P., Tomilin, A., & Scho, H. R. (2009). Post-Translational Regulation of Oct4 Transcriptional Activity, *Ploss One*, 4(2). doi:10.1371
- Quandt, K; Frech, K; Karas, H; Wingender, E; Werner (1995). T. MatInd and matInspector: new fast and versatile tools for detection of consensus matches in nucleotide sequence data. *Nucleic Acids Research*; 23: 4878–4884.
- Rumjanek, V.M.; Lucena, M.; Campos, M.M.; Marques-silva, V.M.; Maia, R. C. (1994). *Multidrug resistance in leukemias: the problem and some approaches to its circumvention. Journal of Brazilian Association for the Advancement of Science*, 46(1/2): 63-69.
- Rumjanek, V.M.; Trindade, G.S.; Wagner-Souza, K.; Meletti-de-Oliveira, M.C.; Marques-Santos, L.F.; Maia, R.C.; Capella, M.A.M. (2001). *Multidrug resistance in tumour cells: characterisation of the multidrug resistant cell line K562-Lucena 1.. An Academia Brasileira de Ciências*, 73(1): 57-69.

- Saxe JP, Tomilin A, Schöler HR, Plath K, Huang J. (2009). Post-translational regulation of Oct4 transcriptional activity. *PLoS One* 4:e4467.
- Scotto, K. W. (2003). Transcriptional regulation of ABC drug transporters. *Oncogene*; 22: 7496–7511.
- Stavrovskaya, A. A.; Stromskaya T. P. (2008). *Transport Proteins of the ABC Family and Multidrug Resistance of Tumor Cells. Biochemistry*; 73: 592-604.
- Tamura, K., e Nei, M. (1993). Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Molecular biology and evolution*, 10(3), 512-26.
- Walker, E., Ohishi, M., Davey, R. E., Zhang, W., Cassar, P. a, Tanaka, T. S., Der, S. D., et al. (2007). Prediction and testing of novel transcriptional networks regulating embryonic stem cell self-renewal and commitment. *Cell stem cell*, 1(1), 71-86. doi:10.1016/j.stem.2007.04.002
- Wang, X., & Dai, J. (2010). Concise review: isoforms of OCT4 contribute to the confusing diversity in stem cell biology. *Stem cells (Dayton, Ohio)*, 28(5), 885-93. doi:10.1002/stem.419
- Wei, F., Schöler, H. R., & Atchison, M. L. (2007). Sumoylation of Oct4 enhances its stability, DNA binding, and transactivation. *The Journal of biological chemistry*, 282(29), 21551-60. doi:10.1074/jbc.M611041200
- Yu J, et al. (2007) Induced pluripotent stem cell lines derived from human somatic cells. *Science* 318:1917–1920.
- Zhang, W., Wang, X., Xiao, Z., Liu, W., Chen, B., & Dai, J. (2010). Mapping of the minimal internal ribosome entry site element in the human embryonic stem cell gene OCT4B mRNA. *Biochemical and biophysical research communications*, 394(3), 750-4. Elsevier Inc. doi:10.1016/j.bbrc.2010.03.064

## **OCT4 mutations in human erythroleukemic cells: implications for multiple drug resistance (MDR) phenotype**

Oliveira, B.R.; Figueiredo, M.A.; Trindade, G.S. and Marins, L.F.

Universidade Federal do Rio Grande – FURG, Instituto de Ciências Biológicas (ICB), Rio Grande, RS – Brazil.

### **Abstract**

The OCT4 transcription factor is a crucial stem cells marker and it has been related to the cancer stem cells (CSCs) concept. More recently, it has also been associated to the multiple drug resistance (MDR) phenotype. Our objective was to test whether OCT4 is linked to the MDR phenotype in erythroleukemic cells derived from LMC-K562 cell line. For this, we performed expression analyses of genes associated to the ATP Binding Cassette (ABC) superfamily, as well as stem cells-related transcription factors. Our first results pointed out the straight correlation between OCT4 and ABC transporters in K562-derivative MDR (Lucena) cells. Sequencing of ABC promoters did not reveal any mutation that could explain the differential expression of OCT4 in Lucena cells. Furthermore, sequencing of the homeobox domain region from the OCT4 gene isolated from both cell lines evinced, for the first time, that this transcription factor is a target of mutations and might be related to the MDR phenotype. The encountered mutations implied in several amino acids substitutions in both cell lines. K562

had seven amino acids substituted (three of them exclusive), while Lucena had 13 substitutions (nine of them exclusive). In addition, an *in silico* search for phosphorylation motifs within the amino acid stretch compared showed that human normal OCT4 has seven potential phosphorylation motifs. However, K562 has lost one phosphorylation motif and Lucena two of them. Molecules presenting different phosphorylation patterns may have its function modified. These findings bring OCT4 as an important target for cancer treatment, especially those resistant to chemotherapy.

**Keywords:** K562, Lucena, gene expression, ABC transporters, transcription factors, phosphorylation motifs

## **Introduction**

The knowledge concerning most part of the known tumors reveals the presence of stem cells sub-populations in it (SCs) (Alison et al., 2012). Within a normal tissue, these cells are responsible for the cellular renewal process due to its great potential of expansion and self-renewal, which is associated with its long life. The SCs present three basic properties. The first concerns the differentiation, which is extremely important for tissues reposition and formation. The differentiation of SCs occurs progressively and originates cell types, which are specific in regards to structure and function, although presenting a short life. The second property of SCs is the ability of self-renewal by originating new SCs, with the same characteristics of their cells of origin, and keeping unaltered the proliferation, expansion and differentiation capacities, allowing the maintenance of SCs storage. The last property of SCs is their capacity of

homeostatic control by modulating and balancing the differentiation and self-renewal processes, according to environmental stimuli, as well as their genetic limitations (Dalerba et al., 2007).

Concerning the fact that SCs are the only long life cells present at most tissues, it is presumable that these may also present higher accumulation of the mutations needed for triggering the carcinogenic process. The accumulated mutations in these cells may lead to loss of DNA repair capacity and apoptosis processes, contributing to tumor formation and survival (HU et al., 2008). Within this context, it raises the concept of cancer stem cells (CSCs). The tumors are composed by heterogenic population of cells supported by CSCs, which also presents the capacity of differentiation or plasticity and self-renewal. Besides, CSCs and SCs share the same capacity of migration, adhesion to many tissues and elevated proliferation rate (KLONISCH et al., 2008). CSCs model was firstly described in cells with chronic myeloid leukemia (CML) presenting a small pool of cell surface markers with the pattern CD34+CD38-, typical of embryonic stem cells (Bonnet e Dick, 1997). These authors conclude that leukemic cells have a primitive origin with side-populations of hematopoietic stem cells, which went under the key events that causes the malignant transformation of leukemic cells and originate CSCs.

Under the light of these new perspectives, the CSCs became a potential target for oncologic treatment. The current methods of cancer treatment are aimed to reduce the tumor mass. However, these methods cannot eliminate all cancerous cells, especially the CSCs, which can regain the original tumoral mass. These events occur because the conventional chemotherapy medicines are effective when the cells are under intensive proliferation during the cellular

cycle. However, studies reveal that normal SCs as well as CSCs remains in G<sub>0</sub>, which is a quiescent state with low mitotic activity (Bertolini et al., 2009; Alison et al., 2012; Nguyen et al., 2012). Therefore, the anti-neoplastic drugs do not exert any effect over CSCs due to this differentiated distinct microenvironment. Although one chemotherapy treatment specifically designed for CSCs would not rapidly reach the tumoral mass, it would still eliminate the renewing focus of the tumor and slowly reduce its volume, which would also drastically reduce its recidive capacity (DALERBA; CHO; CLARKE, 2007). Considering these data, future treatments should focus in this sub-population of tumoral cells and consequently avoid compromising the healthy SCs of the individual, once these are especially important for renewal of tissues damaged by cancer (KLONISCH et al., 2008).

The development of new cancer treatments depends on the identification of specific markers for SCs in the tumoral tissues. The transcription factor OCT4 is a potential target, once it is responsible for keeping the cells in an undifferentiated state. Recently, (HAMMACHI et al., 2012) have shown that the inhibition of this transcription factor allows the triggering of the differentiation process. OCT4 presents a DNA binding domain that enables the recognizing of an octameric nucleotide sequence (ATGCAAAT), which is present at the regulatory regions of many genes involved in cellular differentiation. There were identified 623 possible associations of OCT4 with promoter regions of embryonic stem cells (ESCs) in one study using chromatin immunoprecipitation (ChIP) (WEI; SCHÖLER; ATCHISON, 2007). These results demonstrate that OCT4 acts in the regulation of a large number of genes and plays a central role not only in the embryonic development but also in the maintenance of SCs

storages within adult tissues. OCT4 acts by inhibiting as well as activating genes necessary for pluripotency and it might induce or inhibit differentiation. The reason defining its way of action can be the association with other transcription factors, such as SOX2, NANOG (WALKER et al., 2007) and FOXD3 (LIU; LABOSKY, 2008), or protein post-translational modifications (Plath, et al.; 2009; Wang e Dai, 2010; Zhang et al., 2010).

The OCT4 gene, also known as POU5F1, belongs to the POU family of genes and, in humans, it is localized in position p21.3 of chromosome 6 (WANG; DAI, 2010). This gene codes for a protein with an N-terminal domain (ND), from the amino acid 1 to 133, a POU domain of homeobox type (134-289) and a C-terminal domain (CD), from amino acid 290 to 360 (LIM et al., 2009). There are 3 transcript isoforms (OCT4A, OCT4B e OCT4B1) and, from these, OCT4A is the responsible for stem characteristics (WANG; DAI, 2010). The function of OCT4B isoform is still unknown, although there is speculation about its action also as a stem factor. The OCT4B isoform has additional internal ribosomal entrance sequences (IRES), which allow the translation to three protein isoforms (OCT4B-265, OCT4B-190 e OCT4B-164), and presents OCT4B-190 isoform as protein related to cellular stress (GAO et al., 2010). Besides the already mentioned isoforms, OCT4 also presents pseudogenes, which are characterized by genomic DNA sequences considered extinct relatives of functional genes. Nevertheless, there have already been described genes related to OCT4, that are translated *in vivo*, and might be potentially acting in the regulation of the main OCT4 isoform (LIEDTKE et al., 2007). In addition to its close relation with SCs, OCT4 has also been associated with tumor development. (FANG et al., 2011) have observed a superexpression of



this gene in many types of cancer and associated this characteristic with the cancer stem cells concept. Thus, OCT4 can be considered a specific marker of CSCs, as it was revealed by immunohistochemical marking in several kinds of germinative tumor cells ((LOOIJENGA et al., 2003).

Among the genes controlled by OCT4, there are those coding for proteins of ABC (*ATP binding cassette*) superfamily. This superfamily is composed by approximately 300 transmembranic proteins that transport different compounds out of the cell in an ATP-dependant way (Stavrovskaya and Stromskaya, 2008). These proteins, when present at high concentrations in the membrane, are responsible for the greatest fault concerning the treatment of neoplasias and represent a multiple drug resistance phenotype known as MDR. (BERTOLINI et al., 2009) have shown that OCT4 expression plays a fundamental role in the resistance against the chemotherapy agent cisplatin within lung cancer cells. A more recent study evinces a direct association between CSCs and the MDR phenotype by correlating the superexpression of genes from ABC family and OCT4 from a MDR erythroleukemic lineage resistant to the chemotherapeutic vincristina (VCR) (Marques et al., 2010). However, these same authors observe that the MDR as well as the parental non-MDR lineages present the same CD34+/CD38- profile, highlighting that both lineages present stem characteristics. Therefore, it was not possible to establish a straight relation between OCT4 expression and the MDR phenotype unless there are alternative OCT4 isoforms with distinct capacities of gene regulation, characterizing CSCs sub-populations with differentiated responses against the chemotherapeutic agent.

Within the scenario described above, the present work has tested the hypothesis that might exist specific isoforms of OCT4 related to the MDR phenotype. Therefore, the cellular lineage K562 of LMC was used. This lineage has hematopoietic origin and is equivalent to a human erythroleukemia that presents myeloblasts, as cell type, with lymphoblastic morphology, corresponding to an acute or blastic LMC as established by Lozzio and Lozzio (1975). A K562-MDR derivative cell lineage was established by Rumjanek et al. (1994; 2001), which was selected with VCR and named K562-Lucena1 (Lucena). This lineage already had its MDR phenotype confirmed by many studies (Maia et al., 1996a; 1996b; Marques-Silva, 1996; Orind et al., 1997). The aim of the present work was to study the phenomenon leading to resistance against chemotherapy by comparing the leukemic tumoral lineages K562 and its parental MDR Lucena. Thus, the expression profile from some ABC transporters genes and associated transcription factors were analyzed. Also, the proximal promoter region of ABC genes involved with the MDR phenotype was sequenced, as well as the OCT4 homeobox region. Therefore, it is possible that OCT4 gene alterations, or in its recognition sites of ABC genes regulatory regions, may be tightly related to the previously observed MDR cells response against VCR.

## **Material and Methods**

### *Cells and culture procedures*

The K562 cells (sensitive) and Lucena (MDR) were obtained from the Laboratório de Imunologia Tumoral, no Departamento de Bioquímica Médica da

Universidade Federal do Rio de Janeiro (UFRJ, Brazil). Both lineages were grown in RPMI-1640 medium (GIBCO), inside culture flasks maintained at 37°C, and supplemented with fetal bovine serum (GIBCO), sodium bicarbonate (0.2 g/L), glutamine (0.3 g/L), Hepes (3 g/L), 1% of antibiotic and antimycotic. Besides, the resistant lineage received 60nM VCR (Sigma, Brazil) to maintain the MDR phenotype.

The cellular viability (percentage of viable cells per total number of cells) was evaluated by Trypan blue extrusion method. This method comprises the dilution of cell suspension ( $2 \times 10^6$  cells/mL) with trypan blue solution (0.08%), in the proportion of 1:1, and the addition of the resulting solution into a Neubauer chamber for counting (damaged cells are dyed in blue and viable cells are not dyed).

#### *Total RNA extraction and cDNA synthesis*

Total RNA was extracted by Trizol reagent method, according manufacturer's instructions (Invitrogen, Brazil), using five samples, collected from each cell lineage, presenting a counting of  $3 \times 10^6$  cells per sample. Three flasks with 10 ml RPMI containing about  $300 \times 10^4$  and were disposed in a single falcon tube after centrifugation. Then, a calculation was made to obtain a concentration of  $3 \times 10^6$  cells per sample. RNA from samples was quantified by fluorometry in a Qubit fluorometer (Invitrogen, Brazil) using the Quanti-iT RNA assay kit (Invitrogen, Brazil). RNA integrity was confirmed by electrophoresis in agarose gel (1%) stained with 0.5 µg/mL of ethidium bromide. The RNA was then treated with Dnase1 (Invitrogen, Brazil) and reverse-transcribed (complementary DNA synthesis) using 2 µg of total RNA, following the

manufacturer's instructions of reverse transcriptase (RT High capacity, Invitrogen, Brazil).

#### *Real-Time Quantitative PCR*

The gene expression analysis was realized using a Real-time PCR System 7500 (Applied Biosystems, Brazil). Each PCR reaction was composed by a 10  $\mu$ L mix composed by 5  $\mu$ L of SYBR Green PCR Master Mix (Applied Biosystems, Brazil), 0.2  $\mu$ L of ROX, 0.25  $\mu$ L of each *primer*, 2.7  $\mu$ L of ultrapure water (Invitrogen, Brazil) and 1.6  $\mu$ L of diluted cDNA from the lineages K562 or Lucena. The primers used within this step were designed by using the Primer-Blast tool from the NCBI website (<http://www.ncbi.nlm.nih.gov>). The analyzed genes were from ABC superfamily and transcription factors OCT4, NANOG, SOX2 and FOXD3 after sequences available at GenBank. The gene EF1 $\alpha$  was used as an internal control for data normalization. Information regarding the sequences of primers used for gene expression analysis is show in Table 1. The gene expression data were analyzed by comparison of median $\pm$ SD between the cell lineages employing the software REST (*Relative Expression Software Tool*), according to Pfaffl *et al.* (2002). REST uses the  $2^{-\Delta\Delta C_t}$  Method (Livak & Schmittgen, 2001), followed by a randomization method (Pair Wise Fixed Reallocation Randomization Test) to test the statistical significance of calculated expression ratios To generate standard error estimation, Taylor algorithm is used.

#### *Vincristine (VCR) treatment*

Concerning the gene expression experiments, the multidrug resistant cell line Lucena was not exposed to VCR during 45 days in the first procedure and, differently, it was exposed during four months in the second. Subsequently, the RNA of these cells was extracted in each treatment (with and without VCR), for further gene expression analysis.

#### *Sequencing of gene ABCB1 and ABCG2 promoters*

Genomic DNA from both cell lines was extracted using the kit Axyprep (Axygen Biosciences, Brazil), and an RNase (Invitrogen, Brazil) was used for DNA purification. The quality of the extracted DNA was assessed by electrophoresis in agarose gel (1%). The genes ABCB1 and ABCG2 had their promoters sequenced. Two sets of specific primers were manually designed for each promoter, based on gene sequences available in GenBank, aiming to amplify a region with approximately 1,400 bp (Table 2). The PCR reactions were realized in a thermal cycler under the following conditions: initial denaturation step of 2' at 94°C; 35 cycles of 30'' at 94°C, 30'' at 60°C and 2' at 72°C; and a final extension step of 10' at 72°C. The quality and concentration of the products obtained were assessed by electrophoresis in agarose gel (1%). The Illustra GFX kit (GE Healthcare, Brazil) was used intending to purify the PCR products, which were directly sequenced in a MegaBACE 500 automated sequencer (GE Healthcare, Brazil). Primers and GenBank accession numbers are described in Table 2.

#### *Alignment of the sequenced promoters*

The nucleotide sequences of gene promoters from the two cell lines were compared with sequences available in GenBank using the software BLAST (*Basic Local Alignment Search Tool*). This tool allowed confirming that the isolated promoter was, *de facto*, the equivalent of that present in the human genome. Additionally, each promoter had its sequence from GenBank compared by alignment with those isolated from K562 and Lucena cell lines. The percentage of eventual mutations found in the sequenced promoters was always verified in relation to the sequences available in GenBank.

#### *Mapping OCT4 binding sites in gene promoters*

Promoters were analyzed by the MatInspector program (Quandt *et al.*, 1995). This tool uses a large database of DNA binding sites, which are recognized by transcription factors. A level of similarity is attributed for each site encountered for a determined transcription factor facilitating, therefore, the data filtering and the selection of sites with higher binding potential. OCT4 binding sites were identified not only in the sequenced promoters (ABCB1 and ABCG2), but also in promoters from OCT4 (NC\_000006.11), SOX2 (NC\_000003.11), NANOG (NC\_000012.11) and FOXD3 (NC\_000001.10) available in GenBank.

#### *Sequencing of K562 and Lucena OCT4 coding region*

The PCR reactions from K562 and Lucena cDNAs were performed in a thermal cycler under the following conditions: initial denaturation step of 2' at 94°C; 35 cycles of 30'' at 94°C, 30'' at 60°C and 2' at 72°C; and a final extension step of 10' at 72°C. Specific primers were manually designed on OCT4 5' and 3' UTRs (hOCT4-5UTR: 5'-CTTGGGGCGCCTTCCTTCCC-3');

hOCT4-3UTR: 5'-AGCTCCTCCCCTCCCCCTGT-3'), aiming to amplify the complete OCT4 coding region with approximately 1,150 nucleotides. Subsequently, the presence and concentration of the products obtained were verified by electrophoresis in agarose gel (1%). The Illustra GFX kit (GE Healthcare, Brazil) was used intending to purify the PCR products, which were directly sequenced in a MegaBACE 500 automated sequencer (GE Healthcare, Brazil).

The OCT4 nucleotide sequences from both cell lines were compared with sequences available in GenBank using the software BLAST (*Basic Local Alignment Search Tool*). This tool allowed confirming that the isolated sequences were, indeed, the equivalent of that present in the human transcriptome. Furthermore, K562 and Lucena OCT4 sequences were aligned with human OCT4 sequences from GenBank. The percentage of eventual mutations found was always verified in relation to the sequences available in GenBank. Intending to illustrate the differences among the sequences, it was elaborated a distance tree using Tamura-Nei distance (Tamura and Nei, 1993) and the Neighbor-joining tree build method.

#### *OCT4 amino acids alignment and phosphorylation pattern analysis*

With the objective of identifying possible amino acid substitutions, we performed an alignment of OCT4 amino acid deduced sequences from both K562 and Lucena cell lines, in comparison with human OCT4 available at GenBank. In addition, we analyzed the OCT4 phosphorylation potential motifs using the search tool PhosphoMotif Finder available at the Human Protein Reference Database (<http://www.hprd.org/>) (Peri et al., 2003).

## Results

### *Gene expression analyses*

According to Figures 1a and 1b, cells from Lucena cell line without VCR showed a significant decrease in expression of both ABC and transcription factor genes. Only ABCC11 and NANOG did not present differences in relation to K562 cell line.

After exposure of Lucena cells to VCR during 4 months, it was observed an increased expression in all genes from ABC superfamily analyzed. More specifically, ABCB1 gene had a 30-fold induction, ABCG2 had 9.5-fold, and ABCC2 had 5-fold. The remaining genes also had increased expression at lower levels, except ABCC11 that did not present significant difference (Figure 1c). Concerning transcription factors, SOX2, NANOG and FOX3 did not exhibit significant differences. However, OCT4 presented an elevated expression in the Lucena cell line with a 7-fold induction (Figure 1d).

### *Sequencing of the ABCB1 and ABCG2 gene promoters*

The ABCB1 and ABCG2 proximal promoter regions obtained from K562 and Lucena cells were aligned with the corresponding regions from Genbank. The resulting alignments are showed in Figure 2 (ABCB1) and Figure 3 (ABCG2), where only one mutation was identified in both cells lines at position -384 of the ABCG2 promoter.

### *Mapping OCT4 binding sites in gene promoters*



The use on MatInspector program allowed identifying possible OCT4 binding sites in transcription factors promoters (OCT4, SOX2, NANOG and FOXD3). Table 3 shows the results from these analyses, highlighting the large number of OCT4 potential binding sites in their own promoter (nine) and SOX2 (13), while only seven were detected in NANOG and four in FOXD3. The similarity of the identified potential OCT4 binding sites in relation to the octamer ATGCAAAT, ranged from 0.783 to 0.998. The average similarity for all sites from SOX2 was 0.907, 0.836 for OCT, 0.886 for NANOG, and 0.875 for FOXD3.

#### *Sequencing of K562 and Lucena OCT4 coding region*

There were sequenced 570 nucleotides of the region comprising the homeobox DNA-binding domain from the OCT4 gene in both K562 and Lucena cell lines. The OCT4 sequences were analyzed, and the alignment showed in Figure 4 revealed numerous mutations between the cell lines. When compared to GenBank sequences, K562 showed 11 nucleotide differences. Regarding Lucena, 21 mutations were observed within the same comparison. In addition, there were four exclusive mutations for K562, and 14 for Lucena.

The distance tree of the Figure 5 revealed that the K562 and Lucena cell lines are both distant from the sequences present in GenBank. Although they display similarities in a certain level (seven shared mutations), both of them present mutations that make them distant from each other. The Lucena cell line is the furthest from the original cell lines and it is much more distant than its parental cell line due to a large number of exclusive mutations.

### *OCT4 amino acids alignment and phosphorylation pattern analysis*

The OCT4 deduced amino acid sequences for both K562 and Lucena cells resulted in 189 residues, which were aligned with OCT4 from GenBank in order to identify possible substitutions. The results are showed in Figure 6, reporting an elevated number of changed amino acids in both cell lines. When compared to OCT4 from GenBank, K562 showed seven amino acids substitutions. Three of them were exclusive. Regarding Lucena, it was observed 13 amino acids substitutions, being nine of then exclusive.

Taking in account this outstanding number of amino acid substitutions identified, the next step was searching for potential phosphorylation motifs. The results of this *in silico* analysis are showed in Figure 7. It can be observed that, when compared to OCT4 from GenBank, K562 lost one potential motif, while Lucena has lost two of them.

### **Discussion**

The OCT4 transcription factor is essential for embryonic development and it is required to maintenance of totipotency and pluripotency of stem cells (SCs). Along with other transcription factors, such as SOX2, NANOG and FOXD3, OCT4 exerts its function in pluripotency maintenance by activating genes that keep the cell undifferentiated and repressing other that promote ontogeny (Guo et al., 2002; Kim et al., 2008). Recently, a study involving an erythroleukemic lineage resistant to vincristine (VCR) suggested a relation between the multiple drug resistance phenotype (MDR) and OCT4 (Marques et al., 2010), possibly due the presence of cancer stem cell (CSCs). In the present study we have tested the hypothesis of OCT4/MDR relationship in a more

extensive way, not only by evaluating genes expression, but also by sequencing the proximal promoting region of two important genes involved in the extrusion of chemotherapy agents. Additionally, we have sequenced the homeobox domain region from the OCT4 gene that was isolated from the erythroleukemic lineages K562 (non-MDR) and Lucena (MDR). The obtained results evince a straight relation between the OCT4 gene expression and the MDR phenotype besides revealing, for the first time, a great number of mutations in the coding region of OCT4 gene. Despite the presence of these mutations in the non-MDR lineage, they were two times more often in the resistant lineage, suggesting the presence of a mutated OCT4 gene potentially related to the MDR phenotype.

The first conducted experiment aimed to confirm the relationship between OCT4 gene expression and MDR phenotype in erythroleukemic cells. There were analyzed the OCT4 gene and the associated transcription factors SOX2, NANOG and FOXD3. Concerning the MDR phenotype, there were analyzed the genes coding for the ABC transporters related to chemotherapy agents extrusion ABCB1, ABCG2, ABCC1, ABCC2, ABCC3, ABCC4 and ABCC11 (Scotto, 2003). The results shown in Figure 1 refer to the expression of transcription factors and transporters in the presence and absence of VCR. Concerning this figure, it was evident that the augmentation of transporters expression concomitantly occurs with the increase of OCT4 expression solely in VCR-exposed cells. This observation attests the relation between this transcription factor to the MDR phenotype and raises two hypotheses: (1) there are mutations in ABC gene promoters from MDR line, enabling a higher number of OCT4 connections and, consequently, increasing gene transcription; (2) there are mutations in the own OCT4 gene, from the MDR line, and this

mutated form would be able to find more binding sites in ABC promoters and, therefore, increase gene expression.

Intending to test the first hypothesis, we have isolated and sequenced approximately 1,300 nucleotides of the proximal promoter from two of the most important transporters related to the MDR phenotype: ABCB1 and ABCG2. Indeed, as displayed in Figure 1b, these two genes were the most activated in the presence of VCR in Lucena lineage. Analyses of the obtained sequences for the promoters ABCB1 and ABCG2, which compare the lineages K562 and Lucena to its equivalent sequence available in GenBank, are demonstrated in Figures 2 and 3, respectively. The result evinced one single mutation occurring in ABCG2 promoter of both leukemic lineages and therefore discharged the first hypothesis.

Once rejected the first hypothesis, it is possible that mutated OCT4 types might exist within the leukemic lineages tested. The results concerning the analysis of transcription factors gene expression shown in Figure 1d are evidence that this could be occurring. In this instance, VCR exposure induced an increase of OCT4 solely. Actually, an equivalent increase of SOX2 was expected, once the stoichiometry of these two transcription factors is extremely important to the correct regulation of the related genes (Yu et al., 2007). Considering the tight relationship between OCT4 and SOX2, it is probable that both promoters may be recognized by OCT4. The *in silico* analyses of these promoters, and also the promoters from the other related transcription factors, revealed the degree of interaction among them and their transcriptional self-regulation capacity. The results of these analyses are exhibited in Table 2, where 5,000 nucleotides from OCT4, SOX2, NANOG and FOXD3 promoters

were used for OCT4 binding sites identification. It has been encountered nine sites in the OCT4 promoter, 13 in SOX2, seven in NANOG and four in FOXD3. These observations evinces a close relation between OCT4 and SOX2 principally, in which the first is capable of regulate its own transcription. However, these results do not explain the absence of SOX2 gene transcriptional induction equivalently to OCT4 within the MDR lineage exposed to VCR, even more that the average similarity for all 13 OCT4 binding sites encountered in SOX2 promoter was the highest observed (0.907). One plausible explanation for this issue could be the occurrence of mutations in the OCT4 gene, from the MDR lineage, holding the recognition of the conventional binding sites. The same way that mutations could block the recognition of conventional binding sites, there is the possibility of one mutated OCT4 to recognize new binding sites in promoters that are not the usual targets. This could explain the significant increase observed in the transporters ABCB1 and ABCG2 expression from MDR lineage, since *in silico* analyses of these promoters, performed by Marques et al. (2010), has identified only two binding sites for OCT4 in each one of them.

Aiming to test the second hypothesis, a fragment of 570 nucleotides containing the homeobox domain region from the OCT4 gene was isolated and sequenced for both leukemic lineages tested. This region was chosen because it is responsible for target-promoters recognition by using its capacity of anchoring in DNA binding sites. The obtained sequences were compared with the equivalent sequence available in GenBank and the resulting alignment is shown in Figure 4. Twenty-four mutations were identified among the three compared sequences. The lineage K562, when compared only with the

sequence available in GenBank, has presented 11 mutations, four of them exclusive to the lineage. The Lucena lineage, when compared only with the sequence available in GenBank, has shown 21 mutations with 14 unique to the lineage. These results demonstrate, for the first time, that the OCT4 gene is a target of mutations that could be related to the MDR phenotype. The analysis of distance between the sequences obtained in this study and the human OCT4 sequences available in GenBank is shown in Figure 5. The resulting tree demonstrates that K562 is much closer than Lucena to normal OCT4 sequences. The rising question is: do these mutations imply substitution of amino acids?

Mutations might eventually lead to amino acids substitution with possible function alterations in the resulting protein. Recently, Brumbaugh et al. (2012) have carried out a functional analysis of two amino acids, within the homeobox region of OCT4 gene (T234 and S235), and observed that the loss of the phosphorylation site, induced by mutation, caused an alteration in the regulatory activity of this transcription factor. Intending to evaluate if the mutations observed in the present work imply substitution of amino acids, the deduced amino acid sequences were aligned to the OCT4 protein available in GenBank. The alignment is exhibited in Figure 6 and indicates, when compared with GenBank, that K562 had seven amino acids substituted (three of them exclusive). On the other hand, Lucena presented 13 substitutions with nine of them exclusive. These results reveal that OCT4 of K562 is different from OCT4 of Lucena and both of them are very different from the OCT4 available in GenBank. Undoubtedly, the observed mutations implied in the substitution of

amino acids from both cell lines analyzed. Thus, the new question is: do these substitutions imply in function alteration?

Amino acids substitution can generally affect the protein post-translational processing pattern. Regarding OCT4, some studies suppose that its phosphorylation pattern controls its activity (Brehm et al., 1997; Saxe et al., 2009). The phosphorylation process, in eukaryotes, is one of the main regulatory mechanisms of protein activity in which phosphate groups are transferred by kinase enzymes to specific amino acids such as serine, threonine and tyrosine. Intending to evaluate if the amino acid substitutions observed in the OCT4 of K562 and Lucena causes changes in the phosphorylation pattern from this transcription factor, it was performed an identification analysis of the potential phosphorylation sites. The results of this analysis are shown in Figure 7, in which it can be observed that K562 has lost one phosphorylation site and Lucena two, when compared with OCT4 from GenBank. This observation might have a profound implication in OCT4 activity of K562 but especially in that of Lucena. The manner that these differences may act over the regulation of each OCT4 activity cannot be inferred, although it is likely that they are related to the concept of cancer stem cell and the MDR phenotype.

The results obtained in the present study have certainly demonstrated that the OCT4 gene is related to the MDR phenotype and it is a mutation target within the analyzed erythroleukemic lineages. This fact became more evident when there were not encountered, in this study, any exclusive mutations in the transporters promoters sequenced. The mutation rate calculated for OCT4 of K562 was one per 52 bases and of Lucena was one per 27. These mutations implied into amino acids substitution with consequent alteration in the

phosphorylation pattern of the DNA binding domain from this transcription factor. OCT4 molecules presenting different phosphorylation patterns may control different sets of genes or modify its relation to naturally controlled genes. These findings can open a new field of study regarding cancer treatment, especially those which are resistant to chemotherapy agents, once the characterization of differentiated forms of OCT4 may allow the identification of more specific targets aiming the development of new antineoplastic drugs designed for the combat against chemotherapy resistant tumors.

## References

- Alison, M. R., Lim, S. M. L., & Nicholson, L. J. (2011). Cancer stem cells: problems for therapy? *The Journal of pathology*, 223(2), 147-61. doi:10.1002/path.2793
- Bertolini, G., Roz, L., Perego, P., Tortoreto, M., Fontanella, E., Gatti, L., Pratesi, G., et al. (2009). Highly tumorigenic lung cancer CD133+ cells display stem-like features and are spared by cisplatin treatment. *Proceedings of the National Academy of Sciences of the United States of America*, 106(38), 16281-6. doi:10.1073/pnas.0905653106
- Bonnet D, Dick JE. (1997). Human acute myeloid leukemia is organized as a hierarchy that originates from a primitive hematopoietic cell. *Nature Medicine*, 3:730–37.
- Brehm A, Ohbo K, Schöler H (1997) The carboxy-terminal transactivation domain of Oct-4 acquires cell specificity through the POU domain. *Molecular Cell Biology* 17:154–162.
- Brumbaugh, J., Hou, Z., Russell, J. D., Howden, S. E., Yu, P., Ledvina, A. R., Coon, J. J., et al. (2012). Phosphorylation regulates human OCT4. *Proceedings of the*



- National Academy of Sciences of the United States of America*, 109(19), 7162-8. doi:10.1073/pnas.1203874109.
- Dalerba, P., Cho, R. W., & Clarke, M. F. (2007). Cancer stem cells: models and concepts. *Annual review of medicine*, 58, 267-84. doi:10.1146/annurev.med.58.062105.204854
- Fang, X.-feng, Zhang, W.-yi, Zhao, N., Yu, W., Ding, D., Hong, X., Li, L.-sha, et al. (2011). Genome-wide analysis of OCT4 binding sites in glioblastoma cancer cells. *Journal of Zhejiang University. Science. B*, 12(10), 812-9. doi:10.1631/jzus.B1100059
- Gao, Y., Wang, X., Han, J., Xiao, Z., Chen, B., Su, G., & Dai, J. (2010). The novel OCT4 spliced variant OCT4B1 can generate three protein isoforms by alternative splicing into OCT4B. *Journal of genetics and genomics*, 37(7), 461-5.. doi:10.1016/S1673-8527(09)60065-5
- Guo Y, et al. (2002) The embryonic stem cell transcription factors Oct-4 and FoxD3 interact to regulate endodermal-specific promoter expression. *Proceedings of the National Academy of Sciences of the United States of America*, 99:3663–3667
- Hammachi, F., Morrison, G. M., Sharov, A. A., Livigni, A., Narayan, S., Papapetrou, E. P., O'Malley, J., et al. (2012). Transcriptional Activation by Oct4 Is Sufficient for the Maintenance and Induction of Pluripotency. *Cell Reports*, 1(2), 99-109. doi:10.1016/j.celrep.2011.12.002
- Hu, T., Liu, S., Breiter, D. R., Wang, F., Tang, Y., & Sun, S. (2008). Octamer 4 small interfering RNA results in cancer stem cell-like cell apoptosis. *Cancer research*, 68(16), 6533-40. doi:10.1158/0008-5472.CAN-07-6642
- Kim J, Chu J, Shen X, Wang J, Orkin SH (2008) An extended transcriptional network for pluripotency of embryonic stem cells. *Cell*, 132:1049–1061.
- Klonisch, T., Wiehac, E., Hombach-Klonisch, S., Ande, S. R., Wesselborg, S., Schulze-Osthoff, K., & Los, M. (2008). Cancer stem cell markers in common

- cancers - therapeutic implications. *Trends in molecular medicine*, 14(10), 450-60. doi:10.1016/j.molmed.2008.08.003
- Liedtke, S., Enczmann, J., Waclawczyk, S., Wernet, P., & Kögler, G. (2007). Oct4 and its pseudogenes confuse stem cell research. *Cell stem cell*, 1(4), 364-6. doi:10.1016/j.stem.2007.09.003
- Lim, H.-Y., Do, H.-J., Lee, W.-Y., Kim, D.-K., Seo, H. G., Chung, H.-J., Park, J.-K., et al. (2009). Implication of human OCT4 transactivation domains for self-regulatory transcription. *Biochemical and biophysical research communications*, 385(2), 148-53, doi:10.1016/j.bbrc.2009.05.029
- Livak, K. J., & Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2<sup>(-Delta Delta C(T))</sup> Method. *Methods (San Diego, Calif.)*, 25(4), 402–8. doi:10.1006/meth.2001.1262
- Liu, Y., & Labosky, P. a. (2008). Regulation of embryonic stem cell self-renewal and pluripotency by Foxd3. *Stem cells*, 26(10), 2475-84. doi:10.1634/stemcells.2008-0269
- Looijenga, L. H. J., Stoop, H., Leeuw, H. P. J. C. D., Tumors, C., Roozendaal, K. E. P. V., Zoelen, E. J. J. V., Weber, R. F. A., et al. (2003). POU5F1 ( OCT3 / 4 ) Identifies Cells with Pluripotent Potential in Human Germ Cell Tumors. *Cancer Research* , 63: 2244-2250.
- Lozzio, C.B.; Lozzio, B.B. (1975). Human chronic myelogenous leukemia cell-line with positive Philadelphia chromosome. *Blood*; 45(3): 321-334.
- Maia, R.C.; Silva, E.A.C.; Harab, R.C.; Lucena, M.; Pires, V.; Rumjanek, V.M..(1996a) Sensitivity of vincristine-sensitive K562 and vincristine-resistant K562-Lucena 1 cells to anthracyclines and reversal of multidrug resistance. *Brazilian Journal of Medical Biology Research*, 29(4): 467-472.
- Maia, R.C.; Wagner, K.; Cabral, R.H.; Rumjanek, V.M. (1996b). Heparin reverses Rhodamine 123 extrusion by multidrug resistant cells. *Cancer Letters*; 106: 101-108.

- Marques, D. S.; Sandrini, J. Z.; Boyle, R. T.; Marins, L. F.; Trindade, G. S. (2010). *Relationships between multidrug resistance (MDR) and stem cell markers in human chronic myeloid leukemia cell lines. Leukemia Research*; 34: 757–762.
- Marques-Silva, V.M. (1996). *Efeito de moduladores da diferenciação celular no processo de resistência a múltiplas drogas. Tese de Doutorado, Universidade Federal Fluminense, Niterói, RJ, Brasil.*
- Nguyen, L. V., Vanner, R., Dirks, P., & Eaves, C. J. (2012). Cancer stem cells: an evolving concept. *Nature reviews. Cancer*, 12(2), 133-43. doi:10.1038/nrc3184
- Orind, M.; Wagner-Souza, K.; Maia, R.C.; Rumjanek, V.M. (1997). *Modulation of P-glycoprotein on tumour cells. In: Sotelo JR & Benech JC (Ed.): calcium and cellular metabolism: transport and regulation. New York: Plenum Press*, 72(1): 117-124.
- Peri S, et al. (2003) Development of human protein reference database as an initial platform for approaching systems biology in humans. *Genome Research*, 13:2363–2371.
- Pfaffl, M.W.; Horgan, G.W.; Dempfle, L.,. (2002) Relative expression software tool (REST©) for group-wise comparison and statistical analysis of relative expression results in real-time PCR. *Nucleic Acids Research*, 30: 1–10.
- Plath, K., Huang, J., Saxe, J. P., Tomilin, A., & Scho, H. R. (2009). Post-Translational Regulation of Oct4 Transcriptional Activity, *Ploss One*, 4(2). doi:10.1371
- Quandt, K; Frech, K; Karas, H; Wingender, E; Werner (1995). T. MatInd and matInspector: new fast and versatile tools for detection of consensus matches in nucleotide sequence data. *Nucleic Acids Research*; 23: 4878–4884.
- Rumjanek, V.M.; Lucena, M.; Campos, M.M.; Marques-silva, V.M.; Maia, R. C. (1994). *Multidrug resistance in leukemias: the problem and some approaches to its circumvention. Journal of Brazilian Association for the Advancement of Science*, 46(1/2): 63-69.

- Rumjanek, V.M.; Trindade, G.S.; Wagner-Souza, K.; Meletti-de-Oliveira, M.C.; Marques-Santos, L.F.; Maia, R.C.; Capella, M.A.M. (2001). *Multidrug resistance in tumour cells: characterisation of the multidrug resistant cell line K562-Lucena 1.. An Academia Brasileira de Ciências*, 73(1): 57-69.
- Saxe JP, Tomilin A, Schöler HR, Plath K, Huang J. (2009). Post-translational regulation of Oct4 transcriptional activity. *PLoS One* 4:e4467.
- Scotto, K. W. (2003). Transcriptional regulation of ABC drug transporters. *Oncogene*; 22: 7496–7511.
- Stavrovskaya, A. A.; Stromskaya T. P. (2008). *Transport Proteins of the ABC Family and Multidrug Resistance of Tumor Cells. Biochemistry*; 73: 592-604.
- Tamura, K., e Nei, M. (1993). Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Molecular biology and evolution*, 10(3), 512-26.
- Walker, E., Ohishi, M., Davey, R. E., Zhang, W., Cassar, P. a, Tanaka, T. S., Der, S. D., et al. (2007). Prediction and testing of novel transcriptional networks regulating embryonic stem cell self-renewal and commitment. *Cell stem cell*, 1(1), 71-86. doi:10.1016/j.stem.2007.04.002
- Wang, X., & Dai, J. (2010). Concise review: isoforms of OCT4 contribute to the confusing diversity in stem cell biology. *Stem cells (Dayton, Ohio)*, 28(5), 885-93. doi:10.1002/stem.419
- Wei, F., Schöler, H. R., & Atchison, M. L. (2007). Sumoylation of Oct4 enhances its stability, DNA binding, and transactivation. *The Journal of biological chemistry*, 282(29), 21551-60. doi:10.1074/jbc.M611041200
- Yu J, et al. (2007) Induced pluripotent stem cell lines derived from human somatic cells. *Science* 318:1917–1920.
- Zhang, W., Wang, X., Xiao, Z., Liu, W., Chen, B., & Dai, J. (2010). Mapping of the minimal internal ribosome entry site element in the human embryonic stem cell

gene OCT4B mRNA. *Biochemical and biophysical research communications*,  
394(3), 750-4. Elsevier Inc. doi:10.1016/j.bbrc.2010.03.064

**Table 1.** Primers (forward and reverse) used for gene expression analyses, with their respective accession numbers at GenBank.

| <b>Gene</b>  | <b>Primers (5'-3')</b>   | <b>GenBank<br/>(accession number)</b> |
|--------------|--|---------------------------------------|
| EF1 $\alpha$ | Forward: GCCAGTGGAAACCACGCTGCT<br>Reverse: ATCCTGGAGAGGCAGGGCGCA         | CCDS4980                              |
| ABCB1        | Forward: TCCTCAGTCAAGTTCAGAGTCTTCA<br>Reverse: TCTCCACTTGATGATGTCTCTCACT | CCDS5608                              |
| ABCG2        | Forward: GCAGGGACGAACAATCATCT<br>Reverse: TTCCTGAGGCAATAAGGTG            | CCDS3628                              |
| ABCC1        | Forward: CAGACTGGCAGGGCTACTTC<br>Reverse: CTCACGAAGCAGATGTGGAA           | CCDS42122                             |
| ABCC2        | Forward: AGAGCTGGCCCTTGTACTCA<br>Reverse: TGCCTAGGTAGAGGCTTGGA           | CCDS7484                              |
| ABCC3        | Forward: TGTGCTAGCTGATGGACAGG<br>Reverse: GCAGAGAAAGTTGGCAAAGG           | CCDS32681                             |
| ABCC4        | Forward: AGGCTTCTGTGCGTCATTCT<br>Reverse: CTTGGAGAGGAGTTGCAAGG           | CCDS9474                              |
| ABCC11       | Forward: ACCGCATCCTGGTTATGGGCA<br>Reverse: GGCCATGAGGGCTGCGAACA          | CCDS10732                             |
| OCT4         | Forward: TTCCCATGGCGGGACACCT<br>Reverse: CCCCTGGCCCATCACCTCCA            | CCDS34391                             |
| NANOG        | Forward: GGTGTGACGCAGAAGGCCTCA<br>Reverse: AGTCGGGTTACCAGGCATCC          | CCDS31736                             |
| SOX2         | Forward: AAAAACAGCCCGACCGCGT<br>Reverse: CTCCTGGGCCATCTTGCGCC            | CCDS3239                              |
| FOXD3        | Forward: CTCCTGGGCCATCTTGCGCC<br>Reverse: GGACGGCCACTTCGGGGTTTT          | CCDS624                               |

**Table 2.** *Primers (forward and reverse)* used for isolation and sequencing of ABCB1 and ABCG2 promoters, with their respective accession numbers at GenBank.

| <b>Gene</b> | <b><i>Primers (5'-3')</i></b>   | <b>GenBank<br/>(accession number)</b> |
|-------------|---|---------------------------------------|
| ABCB1       | ABCB1(P1)-FOR: AGCATGCTGAAGAAAGACCACTGC<br>ABCB1(P1)-REV: CCCCACTTGGTCCCATGGACT   | NG_011513                             |
|             | ABCB1(P2)-FOR: CCGTTCCAAGGAGCGCGAGG<br>ABCB1(P2)-REV: TGCTGTGCTCCACTCAGCCA        |                                       |
| ABCG2       | ABCG2(P1)-FOR: AGGAGGACTGTTTGGGAACAGCA<br>ABCG2(P1)-REV: TGCAGGGACAGACTACATGCTCA  | NG_032067                             |
|             | ABCG2(P2)-FOR: TCTCCTGTTCTGACTGCTTGGC<br>ABCG2(P2)-REV: GCTCCTTCAGTAAATGCCTTCAGGT |                                       |

**Table 3.** Mapping of OCT4 binding sites in SOX2, OCT4, NANOG and FOXD3 promoters, evincing position DNA strand and similarity to octamer sequence ATGCAAAT.

| SOX2     |        |            | OCT4     |        |            | NANOG    |        |            | FOXD3    |        |            |
|----------|--------|------------|----------|--------|------------|----------|--------|------------|----------|--------|------------|
| Position | Strand | Similarity | Position | Strand | Similarity | Position | Strand | Similarity | Position | Strand | Similarity |
| 795      | +      | 0,887      | 24       | +      | 0,783      | 1133     | +      | 0,833      | 296      | -      | 0,984      |
| 831      | -      | 0,945      | 555      | -      | 0,845      | 1321     | -      | 0,816      | 941      | +      | 0,805      |
| 837      | -      | 0,951      | 743      | -      | 0,951      | 1419     | +      | 0,829      | 1509     | -      | 0,828      |
| 865      | +      | 0,806      | 1089     | +      | 0,809      | 2580     | +      | 0,857      | 3664     | -      | 0,884      |
| 1140     | -      | 0,812      | 1090     | -      | 0,79       | 3154     | -      | 0,927      |          |        |            |
| 1482     | -      | 0,975      | 2506     | +      | 0,87       | 3501     | +      | 0,941      |          |        |            |
| 1564     | -      | 0,985      | 2709     | -      | 0,807      | 4673     | +      | 0,998      |          |        |            |
| 1567     | +      | 0,906      | 3124     | -      | 0,853      |          |        |            |          |        |            |
| 1759     | -      | 0,901      | 3155     | +      | 0,813      |          |        |            |          |        |            |
| 2000     | -      | 0,914      |          |        |            |          |        |            |          |        |            |
| 2071     | -      | 0,834      |          |        |            |          |        |            |          |        |            |
| 2527     | +      | 0,941      |          |        |            |          |        |            |          |        |            |
| 4466     | -      | 0,935      |          |        |            |          |        |            |          |        |            |



## Figure legends

**Figure 1.** Gene expression of seven transporters from ABC superfamily and transcription factors OCT4, SOX2, NANOG and FOXD3 in Lucena (MDR) cells, comparing to its parental non-MDR K562. (a) and (b) Lucena cells without vincristina (VCR) for 45 days; (c) and (d) Lucena cells treated with VCR for four months. The expression of all genes in K562 line was considered = 1, represented by the horizontal line parallel to the X-axis. Asterisks correspond to statistically significant differences ( $p < 0.05$ ). (descrever o teste do REST)

**Figure 2.** Alignment of 1,158 nucleotides from proximal promoter region of ABCB1 gene, comparing K562 and Lucena cell lines with the corresponding sequence from human genome available at GenBank.

**Figure 3.** Alignment of 1,146 nucleotides from proximal promoter region of ABCG2 gene, comparing K562 and Lucena cell lines with the corresponding sequence from human genome available at GenBank.

**Figure 4.** Alignment of 570 nucleotides from OCT4 coding region, comparing K562 and Lucena cell lines with the corresponding sequence from OCT4 human gene available at GenBank.

**Figure 5.** Distance tree among human OCT4 sequences available at GenBank and OCT4 from K562 and Lucena cell lines. The number of shared and exclusive mutations is given at branches.

**Figure 6.** Alignment of 189 deduced amino acids from OCT4, comparing K562 and Lucena cell lines with the corresponding sequence from human OCT4 available at GenBank.

**Figure 7.** Phosphorylation motifs encountered at OCT4 from K562 and Lucena cell lines, comparing with motifs present at human OCT4 available at GenBank. The homeobox domain is delimited by the black rectangle. The circled phosphate groups represent differences in the phosphorylation pattern among compared OCT4 sequences. The associated tables detail the localization and types of phosphorylation motifs detected by the Phospho-motif Finder (<http://www.hprd.org/>).

Figure 1

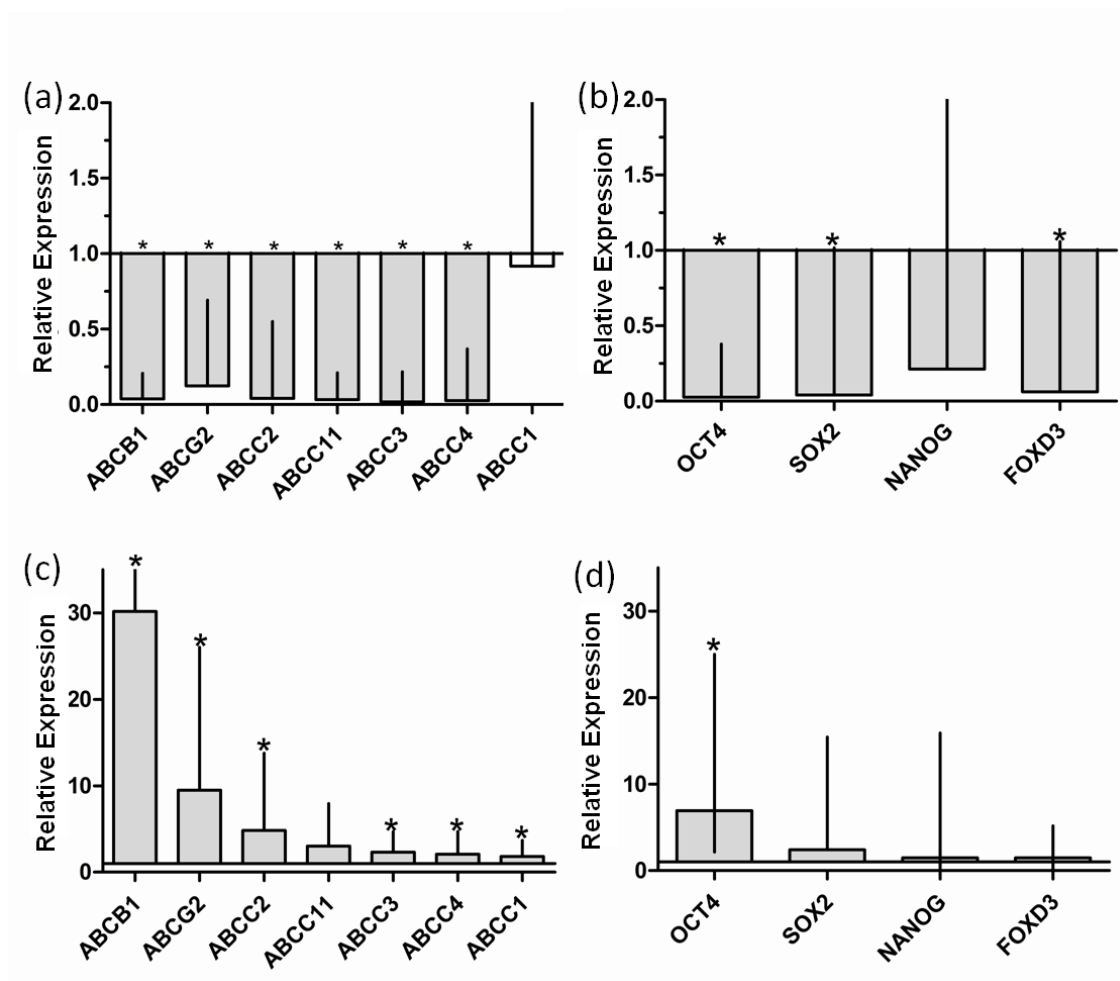


Figure 2

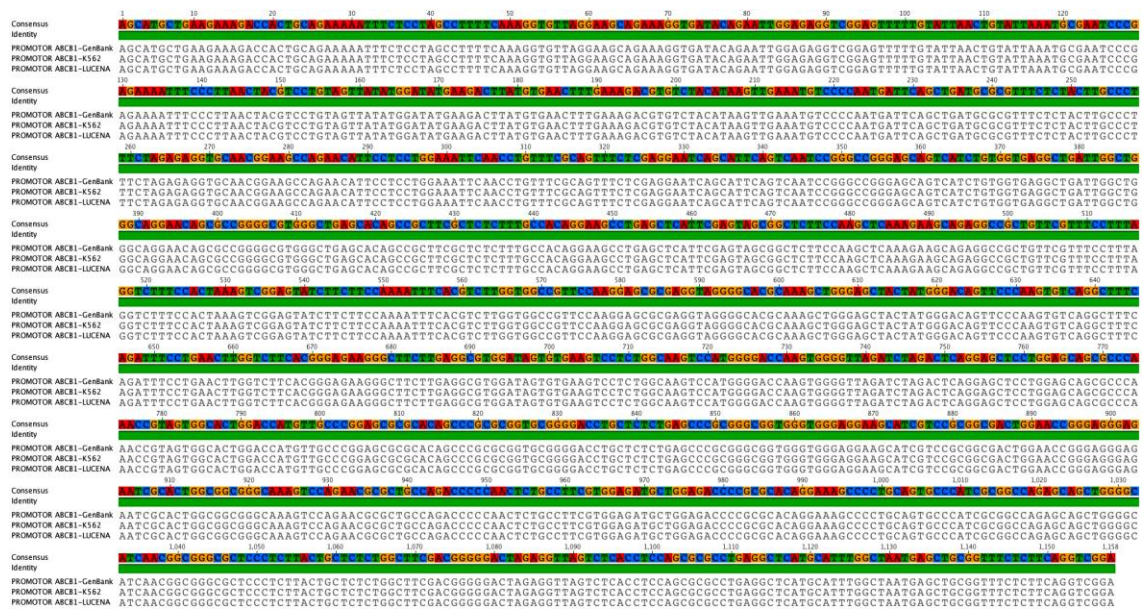






Figure 5

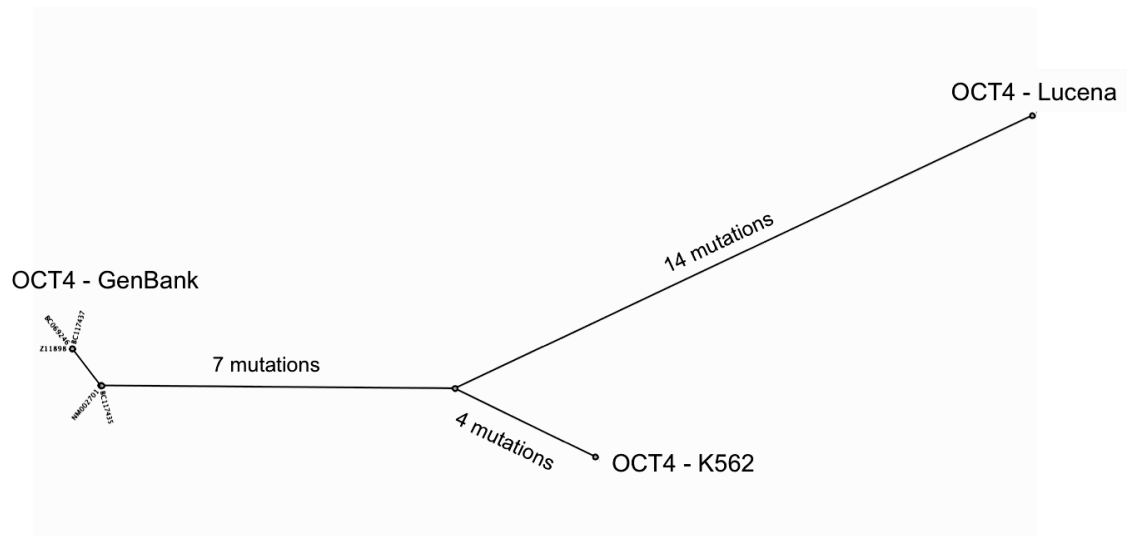


Figure 6

|              |     |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |     |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |  |  |
|--------------|-----|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|-----|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|--|--|
| Identity     | 1   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   | 10  |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |  |  |
| OCT4 GenBank | L   | T | L | G | V | L | F | G | K | V | F | S | O | T | T | I | C | R | R | L | T   | L | G | V | L | F | G | K | V | F | S | O | T | T | I | C | R | R |  |  |
| OCT4-K562    | L   | T | L | G | V | L | F | G | K | V | F | S | O | T | T | I | C | R | R | L | T   | L | G | V | L | F | G | K | V | F | S | O | T | T | I | C | R | R |  |  |
| OCT4-LUCENA  | L   | T | L | G | V | L | F | G | K | V | F | S | O | T | T | I | C | R | R | L | T   | L | G | V | L | F | G | K | V | F | S | O | T | T | I | C | R | R |  |  |
| Identity     | 20  |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   | 30  |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |  |  |
| OCT4 GenBank | F   | E | A | L | Q | L | S | F | K | N | M | C | K | L | R | P | L | L | F | E | A   | L | Q | L | S | F | K | N | M | C | K | L | R | P | L | L |   |   |  |  |
| OCT4-K562    | F   | E | A | L | Q | L | S | F | K | N | M | C | K | L | R | P | L | L | F | E | A   | L | Q | L | S | F | K | N | M | C | K | L | R | P | L | L |   |   |  |  |
| OCT4-LUCENA  | F   | E | A | L | Q | L | S | F | K | N | M | C | K | L | R | P | L | L | F | E | A   | L | Q | L | S | F | K | N | M | C | K | L | R | P | L | L |   |   |  |  |
| Identity     | 40  |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   | 50  |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |  |  |
| OCT4 GenBank | Q   | K | W | V | E | E | A | D | N | N | E | N | L | O | E | I | C | K | Q | K | W   | V | E | E | A | D | N | N | E | N | L | O | E | I | C | K |   |   |  |  |
| OCT4-K562    | Q   | K | W | V | E | E | A | D | N | N | E | N | L | O | E | I | C | K | Q | K | W   | V | E | E | A | D | N | N | E | N | L | O | E | I | C | K |   |   |  |  |
| OCT4-LUCENA  | Q   | K | W | V | E | E | A | D | N | N | E | N | L | O | E | I | C | K | Q | K | W   | V | E | E | A | D | N | N | E | N | L | O | E | I | C | K |   |   |  |  |
| Identity     | 60  |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   | 70  |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |  |  |
| OCT4 GenBank | A   | E | T | L | V | Q | A | R | K | R | K | R | R | T | S | I | E | N | R | A | E   | T | L | V | Q | A | R | K | R | K | R | R | T | S | I | E | N | R |  |  |
| OCT4-K562    | A   | E | T | L | V | Q | A | R | K | R | K | R | R | T | S | I | E | N | R | A | E   | T | L | V | Q | A | R | K | R | K | R | R | T | S | I | E | N | R |  |  |
| OCT4-LUCENA  | A   | E | T | L | V | Q | A | R | K | R | K | R | R | T | S | I | E | N | R | A | E   | T | L | V | Q | A | R | K | R | K | R | R | T | S | I | E | N | R |  |  |
| Identity     | 80  |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   | 90  |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |  |  |
| OCT4 GenBank | V   | R | G | N | L | E | N | L | F | L | Q | C | P | K | P | T | L | O | V | R | G   | N | L | E | N | L | F | L | Q | C | P | K | P | T | L | O |   |   |  |  |
| OCT4-K562    | V   | R | G | N | L | E | N | L | F | L | Q | C | P | K | P | T | L | O | V | R | G   | N | L | E | N | L | F | L | Q | C | P | K | P | T | L | O |   |   |  |  |
| OCT4-LUCENA  | V   | R | G | N | L | E | N | L | F | L | Q | C | P | K | P | T | L | O | V | R | G   | N | L | E | N | L | F | L | Q | C | P | K | P | T | L | O |   |   |  |  |
| Identity     | 100 |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   | 110 |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |  |  |
| OCT4 GenBank | Q   | I | S | H | I | A | Q | Q | L | G | L | E | K | D | V | V | R | V | Q | I | S   | H | I | A | Q | Q | L | G | L | E | K | D | V | V | R | V |   |   |  |  |
| OCT4-K562    | Q   | I | S | H | I | A | Q | Q | L | G | L | E | K | D | V | V | R | V | Q | I | S   | H | I | A | Q | Q | L | G | L | E | K | D | V | V | R | V |   |   |  |  |
| OCT4-LUCENA  | Q   | I | S | H | I | A | Q | Q | L | G | L | E | K | D | V | V | R | V | Q | I | S   | H | I | A | Q | Q | L | G | L | E | K | D | V | V | R | V |   |   |  |  |
| Identity     | 120 |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   | 130 |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |  |  |
| OCT4 GenBank | W   | F | C | N | R | R | Q | K | G | K | R | S | S | S | D | Y | A | Q | W | F | C   | N | R | R | Q | K | G | K | R | S | S | S | D | Y | A | Q |   |   |  |  |
| OCT4-K562    | W   | F | C | N | R | R | Q | K | G | K | R | S | S | S | D | Y | A | Q | W | F | C   | N | R | R | Q | K | G | K | R | S | S | S | D | Y | A | Q |   |   |  |  |
| OCT4-LUCENA  | W   | F | C | N | R | R | Q | K | G | K | R | S | S | S | D | Y | A | Q | W | F | C   | N | R | R | Q | K | G | K | R | S | S | S | D | Y | A | Q |   |   |  |  |
| Identity     | 140 |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   | 150 |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |  |  |
| OCT4 GenBank | R   | E | D | F | E | A | A | G | S | P | F | S | G | G | P | V | S | F | R | E | D   | F | E | A | A | G | S | P | F | S | G | G | P | V | S | F |   |   |  |  |
| OCT4-K562    | R   | E | D | F | E | A | A | G | S | P | F | S | G | G | P | V | S | F | R | E | D   | F | E | A | A | G | S | P | F | S | G | G | P | V | S | F |   |   |  |  |
| OCT4-LUCENA  | R   | E | D | F | E | A | A | G | S | P | F | S | G | G | P | V | S | F | R | E | D   | F | E | A | A | G | S | P | F | S | G | G | P | V | S | F |   |   |  |  |
| Identity     | 160 |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   | 170 |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |  |  |
| OCT4 GenBank | P   | L | A | P | G | P | H | F | G | T | P | G | Y | G | S | P | H | F | P | L | A   | P | G | P | H | F | G | T | P | G | Y | G | S | P | H | F |   |   |  |  |
| OCT4-K562    | P   | L | A | P | G | P | H | F | G | T | P | G | Y | G | S | P | H | F | P | L | A   | P | G | P | H | F | G | T | P | G | Y | G | S | P | H | F |   |   |  |  |
| OCT4-LUCENA  | P   | L | A | P | G | P | H | F | G | T | P | G | Y | G | S | P | H | F | P | L | A   | P | G | P | H | F | G | T | P | G | Y | G | S | P | H | F |   |   |  |  |
| Identity     | 180 |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   | 190 |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |  |  |
| OCT4 GenBank | T   | A | L | Y | S | S | V | P | F | P | E | G | E | A | F | P | P | V | T | A | L   | Y | S | S | V | P | F | P | E | G | E | A | F | P | P | V |   |   |  |  |
| OCT4-K562    | T   | A | L | Y | S | S | V | P | F | P | E | G | E | A | F | P | P | V | T | A | L   | Y | S | S | V | P | F | P | E | G | E | A | F | P | P | V |   |   |  |  |
| OCT4-LUCENA  | T   | A | L | Y | S | S | V | P | F | P | E | G | E | A | F | P | P | V | T | A | L   | Y | S | S | V | P | F | P | E | G | E | A | F | P | P | V |   |   |  |  |
| Identity     | 200 |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   | 210 |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |  |  |
| OCT4 GenBank | S   | V | T | T | L | G | S | P | M |   |   |   |   |   |   |   |   |   | S | V | T   | T | L | G | S | P | M |   |   |   |   |   |   |   |   |   |   |   |  |  |
| OCT4-K562    | S   | V | T | T | L | G | S | P | M |   |   |   |   |   |   |   |   |   | S | V | T   | T | L | G | S | P | M |   |   |   |   |   |   |   |   |   |   |   |  |  |
| OCT4-LUCENA  | S   | V | T | T | L | G | S | P | M |   |   |   |   |   |   |   |   |   | S | V | T   | T | L | G | S | P | M |   |   |   |   |   |   |   |   |   |   |   |  |  |



**Figure 7**

| Protein Position | Sequence in the protein | Motifs (phosphorylated residues in red) | Phospho-motif types                     |
|------------------|-------------------------|---|---|
| 1 119 - 122      | RSSS                    | RXX <sup>o</sup> S                      | 14-3-3 domain binding motif             |
| 2 120 - 122      | SSS                     | S(pS/pT)X                               | MDC1 BRCT/ Pk1 PBD domain binding motif |
| 3 135 - 136      | SP                      | (pS/pT)P                                | WW domain binding motif                 |
| 4 154 - 155      | TP                      | (pS/pT)P                                | WW domain binding motif                 |
| 5 159 - 160      | SP                      | (pS/pT)P                                | WW domain binding motif                 |
| 6 167 - 169      | SSV                     | S(pS/pT)X                               | MDC1 BRCT/ Pk1 PBD domain binding motif |
| 7 187 - 188      | SP                      | (pS/pT)P                                | WW domain binding motif                 |



| Protein Position | Sequence in the protein | Motifs (phosphorylated residues in red) | Phospho-motif types                     |
|------------------|-------------------------|---|---|
| 1 119 - 122      | RSSS                    | RXX <sup>o</sup> S                      | 14-3-3 domain binding motif             |
| 2 120 - 122      | SSS                     | S(pS/pT)X                               | MDC1 BRCT/ Pk1 PBD domain binding motif |
| 3 154 - 155      | TP                      | (pS/pT)P                                | WW domain binding motif                 |
| 4 159 - 160      | SP                      | (pS/pT)P                                | WW domain binding motif                 |
| 5 167 - 169      | SSV                     | S(pS/pT)X                               | MDC1 BRCT/ Pk1 PBD domain binding motif |
| 6 187 - 188      | SP                      | (pS/pT)P                                | WW domain binding motif                 |



| Protein Position | Sequence in the protein | Motifs (phosphorylated residues in red) | Phospho-motif types                     |
|------------------|-------------------------|---|---|
| 1 120 - 122      | SSS                     | S(pS/pT)X                               | MDC1 BRCT/ Pk1 PBD domain binding motif |
| 2 154 - 155      | TP                      | (pS/pT)P                                | WW domain binding motif                 |
| 3 159 - 160      | SP                      | (pS/pT)P                                | WW domain binding motif                 |
| 4 167 - 169      | SSV                     | S(pS/pT)X                               | MDC1 BRCT/ Pk1 PBD domain binding motif |
| 5 187 - 188      | SP                      | (pS/pT)P                                | WW domain binding motif                 |

