Análise da via de regulação gênica por ácido retinóico:
uma abordagem por bioinformática e biologia estrutural

Tiago José Paschoal Sobreira

Tese apresentada ao Programa de Pós-Graduação Interunidades em Bioinformática, Universidade de São Paulo, para a obtenção de título de Doutor em Bioinformática.

Orientadores:

Prof. Dr. Paulo Sergio Lopes de Oliveira

Prof. Dr. Carlos Alberto de Bragança Pereira

São Paulo

2008

Aos meus pais

A meu pai por ensinar a trabalhar e
sempre fazer o mais bem feito possível.

A minha mãe por ensinar a viver e
sempre viver da melhor maneira possível.

AGRADECIMENTOS

Ao Dr. Paulo Sérgio Lopes de Oliveira pela orientação, formação profissional e amizade.

Ao Prof. Dr. Carlos Alberto de Bragança Pereira pela co-orientação.

Ao Dr. José Xavier-Neto pela constante ajuda e grande contribuição nesse trabalho.

Ao Prof. Dr. José Eduardo Krieger por permitir o desenvolvimento desta tese no Laboratório de Genética e Cardiologia Molecular do Instituto do Coração, Hospital das Clínicas, Faculdade de Medicina, Universidade de São Paulo.

Ao Prof. Dr. Arthur Gruber pela formação profissional e amizade.

Ao Prof. Dr. Fernando Portella de Luna Marques por disponibilizar o cluster de computadores do Laboratório de Sistemática Molecular do Instituto de Biociências da Universidade de São Paulo (IB-USP).

Ao Programa de Pós-Graduação Interunidades em Bioinformática pela oportunidade de realização do curso de doutorado.

À Sra. Patrícia, secretária do PPG em Bioinformática, pelo excelente apoio.

À Coordenadoria (CAPES) pela concessão da bolsa de doutorado.

As funcionárias do Laboratório de cardiologia molecular Isabel Cristina e Silvana pelas constantes ajudas.

Aos meus irmãos e cunhadas pela eterna amizade.

Aos meus sobrinhos por alegrar o meu caminho.

Ao Fabio Passetti pela grande ajuda no desenvolvimento deste trabalho e pela amizade e companheirismo.

Ao Marcio Almeida pelos excelentes almoços e pelas grandes discussões científicas e não científicas.

Ao Tiago Pereira pelos excelentes lanches da tarde com soluções de várias dúvidas científicas e não científicas e pelas contribuições nesse trabalho.

A Manuela Gimenes pelas contribuições nesse trabalho e pela grande amizade.

A Daniela Castro pela amizade, carinho, constante companheirismo e por ter aparecido na minha vida!!!

A Elizabeth Gorgone pela amizade, carinho, constante companheirismo.

A Daniel Rabanea pela grande amizade, constante companheirismo.

Aos grandes amigos Daneil Carneiro, Célia Boyadjian, Almir Picanço, Carolina Gois, Cristiane Minussi, Danilo Coelho, Juliana Velasco, Ernesto Nakayasu, Leandro Tambosi, Guilherme Ide, Therry Faria Lima, Mariana Farias, Priscila Rovina e Joana DeCapitani companheiros constantes e queridos da vida.

A Camila Holpert, Denise Gonçalves, Karina Noçais e Renata Estender pela eterna amizade.

RESUMO

Tese (Doutorado) – Programa de Pós-Graduação Interunidades em Bioinformática, Universidade de São Paulo, São Paulo, 2008.

As vias de sinalização celular por meio de moléculas são um dos principais meios de controle funcional de um organismo. O entendimento das funções de moléculas sinalizadoras facilita a compreensão das vias metabólicas de um organismo, assim possibilitando uma melhor compreensão de vários eventos biológicos e também de várias doenças. A sinalização pelo ácido retinóico (AR), e seus derivados, é responsável pelo controle de várias funções, por exemplo: crescimento celular, diferenciação celular, formação da retina, desenvolvimento cardíaco e também relacionado a várias patologias como diabetes, obesidades, cânceres, e doenças cardiovasculares.

A ação do ácido retinóico é controlada em dois níveis: no metabolismo de síntese/degradação e na sua utilização na sinalização para a expressão gênica.

A maquinaria que controla o metabolismo inclui as enzimas de síntese do AR (aldeído desidrogenase ALDH) e as enzimas de degradação do AR (Cyp26), que controlam a distribuição espaço-temporal do AR durante a embriogênese. As ALDHs são enzimas NAD(P)⁺ dependentes, que oxidam uma ampla gama de aldeídos para os seus correspondentes ácidos carboxílicos, sendo ALDH1A2 a principal enzima na transformação de retinal em ácido retinóico.

A maquinaria da sinalização celular por AR contém os receptores nucleares controlados por AR (RARs) que estão envolvidos com o controle da transcrição gênica. Os mecanismos de controle de expressão mais comuns são os que ocorrem na fase transcrricional. Um desses mecanismos envolve proteínas que se ligam às regiões promotoras de transcrição, representadas por trechos de DNA que geralmente estão localizados próximo à região de início da transcrição, mas que também podem estar a centenas ou até milhares de pares de bases desse início. Essas proteínas modulam a maquinaria transcrricional, podendo ativá-la ou inibi-la.

A associação de várias técnicas como a biologia molecular, bioinformática, filogenia, análises estruturais de biomoléculas, mecânica molecular e métodos termodinâmicos tem se mostrado uma poderosa abordagem para compreensão de sistemas biológicos simplificando e agilizando o desenvolvimento do conhecimento científico. Nessa direção, esse estudo desenvolveu duas análises: a primeira estudando a evolução das funções das enzimas ALDH, utilizando-se de técnicas de genômica combinatória, filogenia, bioinformática, estrutura de biomoléculas e de biologia do desenvolvimento, tentando compreender o modo como as ALDHs, que apresentam as seqüências de aminoácidos bastante similares, puderam divergir para gerar funções diversas como a destoxificação e a sinalização. Para este estudo foram analisados os genomas de 487 organismos em busca de seqüências de ALDHs e também o genoma do organismo modelo *Branchiostoma floridae*. Foram obtidas 190 seqüências que foram utilizadas em uma análise filogenética para tentar compreender a função primordial e também para definir grupos de aminoácidos candidatos a marcadores das diferentes famílias de ALDHs. Essas 190 seqüências também foram modeladas estruturalmente e analisada a forma e o volume do canal onde se aloja o aldeído a ser oxidado. A partir dessas informações foi possível prever que as ALDHs passaram das funções ancestrais de controle do padrão corporal para algo mais abrangente como funções protetoras.

A segunda análise, utilizando-se das estruturas tridimensionais dos fatores de transcrição ligados ao DNA em diferentes posições e submetendo esses complexos a processos de mecânica molecular, cálculos termodinâmicos e análises das ligações de hidrogênio para tentar prever os mais prováveis sítios de interação entre os receptores e o DNA. O modelo escolhido para essa análise foram os fatores de transcrição regulados por ácido retinóico o RAR e RXR utilizando a região promotora do gene RARE- β 2 para avaliar as mais prováveis regiões de ligação desses fatores. Para esse estudo foram construídos 71 complexos proteína-DNA que foram submetidos a processos de mecânica molecular e cálculos termodinâmicos. A partir dessas informações foi possível prever uma região de maior afinidade entre o fator de transcrição e o DNA. As análises de ligações de hidrogênio possibilitaram definir exatamente a região de interação entre os fatores de transcrição e o DNA, e também descrever as interações moleculares responsáveis pela especificidade da interação.

Palavras-chave: Bioinformática. Sinalização celular. Modelagem molecular por homologia. Fatores de transcrição. Mecânica molecular. Ácido retinóico.