

Universidade Federal De Pernambuco
Laboratório De Imunopatologia Keizo Asami
Programa De Pós-Graduação Em Biologia Aplicada À Saúde

MARIA AMÉLIA CARLOS SOUTO MAIOR BORBA

**ANÁLISE *IN SILICO* DOS POLIMORFISMOS DA CYP2D6 E SUA INFLUÊNCIA
NA FARMACOGENÔMICA DO CÂNCER DE MAMA**

Recife
2016

MARIA AMÉLIA CARLOS SOUTO MAIOR BORBA

**ANÁLISE *IN SILICO* DOS POLIMORFISMOS DA CYP2D6 E SUA INFLUÊNCIA
NA FARMACOGENÔMICA DO CÂNCER DE MAMA**

Dissertação de Mestrado apresentada ao
Programa de Pós Graduação em Biologia
Aplicada à Saúde do Laboratório de
Imunopatologia Keizo Asami – LIKA/UFPE,
como requisito para a obtenção do título de
mestre.

Orientadora: Danyelly Bruneska Gondim
Martins

Recife
2016

Catálogo na Fonte:
Bibliotecário Bruno Márcio Gouveia, CRB-4/1788

Borba, Maria Amélia Carlos Souto Maior

Análise *in silico* dos polimorfismos da CYP2D6 e sua influência na farmacogenômica do câncer de mama / Maria Amélia Carlos Souto Maior Borba. – Recife: O Autor, 2016.

54 f.: il.

Orientadora: Danyelly Bruneska Gondim Martins

Dissertação (mestrado) – Universidade Federal de Pernambuco. Centro de Biociências. Programa de Pós-graduação em Biologia Aplicada à Saúde, 2016.

Inclui referências

1. Mamas – Câncer 2. Bioinformática 3. Polimorfismo (Genética) I. Martins, Danyelly Bruneska Gondim (orient.) II. Título.

616.99449
172

CDD (22.ed.)

UFPE/CCB-2017-

MARIA AMÉLIA CARLOS SOUTO MAIOR BORBA

**ANÁLISE *IN SILICO* DOS POLIMORFISMOS DA CYP2D6 E SUA
INFLUÊNCIA NA FARMACOGENÔMICA DO CÂNCER DE MAMA**

Dissertação de Mestrado apresentada
ao Programa de Pós Graduação em
Biologia Aplicada à Saúde do
Laboratório de Imunopatologia Keizo
Asami – LIKA/UFPE, como requisito
para a obtenção do título de mestre.

Orientadora: Danyelly Bruneska
Gondim Martins

Aprovada em 15 de abril de 2016

COMISSÃO EXAMINADORA

Prof^a. Dr^a. Danyelly Bruneska Gondim Martins/UFPE

Prof^a. Dr^a. Cíntia Renata Costa Rocha/UFPE

Prof. João Bosco de Oliveira Filho/Genomika Diagnósticos

Prof^a. Dr^a. Rosângela Ferreira Frade de Araújo/UFPE

Prof. Dr. Eduardo Isidoro Carneiro Beltrão/UFPE

Ao meu bisavô Sebastião Souto Maior, de quem herdei a
ciência e ao meu avô João Borges, pelo orgulho que tinha
de mim; em nome dos quais a toda minha família,

Dedico

AGRADECIMENTOS

A Deus e meu anjo da guarda, pelas oportunidades e pelo cuidado em me fazer a cada dia, uma pessoa um pouco melhor.

A minha orientadora, Professora Danyelly por ter me acolhido e me concedido participar deste projeto, por ter me acompanhado passo a passo neste processo de crescimento e por além de acreditar nos meus sonhos, ter sonhado junto comigo.

A Henrique Castelletti pela paciência e disponibilidade para os desafios que a bioinformática me trouxe.

Ao professor José Luiz por ter me recebido no LIKA há cinco anos atrás, quando dei início as atividades de iniciação científica no grupo de biossensores.

Aos colegas do grupo, especialmente aos do grupo de *ProspecInfo*, pela colaboração e por terem feito meus dias mais leves, me proporcionando boas risadas mesmo nos momentos mais desesperadores.

Ao Programa de Pós-Graduação em Biologia Aplicada à Saúde pela dedicação exercida na formação de recursos humanos comprometidos.

Ao LIKA e todos que o fazem funcionar diariamente, pela infraestrutura e por ter constituído aqui minha segunda casa, e logo, minha segunda família.

À FACEPE que fomentou este projeto de pesquisa.

Aos meus pais que sempre me apoiaram em todas as minhas decisões, mesmo quando isto custou minha ausência.

Ao meu namorado Renato pelo apoio incondicional e ininterrupto.

Todo o caminho da gente é resvaloso.
Mas também, cair não prejudica demais
A gente levanta, a gente sobe, a gente volta!
O correr da vida embrulha tudo, a vida é assim:
Esquenta e esfria, aperta e daí afrouxa,
Sossega e depois desinquieta.
O que ela quer da gente é coragem.
Ser capaz de ficar alegre e mais alegre no meio da alegria,
E ainda mais alegre no meio da tristeza.

João Guimarães Rosa, Veredas.

RESUMO

O câncer de mama é o tipo que possui a maior taxa de mortalidade entre mulheres de todo o mundo. Ele possui diferentes características moleculares, classificando-se em: Luminais A e B, HER2 e basal-like. O tamoxifeno, tratamento clássico dos tumores mamários luminais, é uma pró-droga que precisa ser metabolizada por enzimas hepáticas, especialmente a CYP2D6, para exercer suas funções terapêuticas. No entanto, cerca de 40% das pacientes não respondem adequadamente ao tratamento com tamoxifeno e existem evidências que apontam a perda de função da CYP2D6 como causa das respostas indesejáveis do tamoxifeno, embora os dados sejam conflitantes. A CYP2D6 é uma enzima do P450, altamente polimórfica, fenotipicamente classificada de acordo com a atividade em metabolizadores pobres, intermediários, rápidos e extensivos. O objetivo deste trabalho foi aplicar ferramentas de bioinformática para avaliar o impacto das mutações de sentido trocado na estrutura e função de alelos metabolizadores pobres da CYP2D6. A base de dados CYPalleles foi consultada para catalogar os alelos pobre metabolizadores da CYP2D6 e suas respectivas mutações e selecionar alelos que contêm exclusivamente mutações de sentido trocado. Todos os polimorfismos presentes nos alelos selecionados, CYP2D6*7 e CYP2D6*14A, foram submetidos a análise através de sete algoritmos de predição de impacto estrutural e/ou funcional na proteína. Subsequentemente foram construídos modelos 3D baseados na estrutura terciária alterada pelas mutações. Os resultados evidenciaram que a única mutação presente no alelo CYP2D6*7, a H324P, compromete a dinâmica do sítio ativo devido a sua localização entre duas alfas hélices, resultando em alto impacto funcional. Por outro lado, o alelo CYP2D6*14A possui quatro mutações, os polimorfismos R296C e S486T estão presentes em vários alelos da CYP2D6, e as análises *in silico* não apontam significantes efeitos deletérios. A mutação P34S também está presente em vários alelos, mas foi predita como deletéria devido a considerável efeito estabilizante pelo surgimento de uma nova ponte de hidrogênio. Por outro lado a G169R caracteriza genotipicamente o alelo CYP2D6*14 e leva a aumenta a acessibilidade ao solvente com severa desestabilização da enzima, pois está presente numa região de interface entre as cadeias monoméricas. As análises *in silico* evidenciaram efeitos deletérios das mutações de sentido trocado na estrutura e função da CYP2D6, o que pode explicar a resistência ao tratamento com tamoxifeno.

Palavras-chave: Bioinformática. Câncer de mama. CYP2D6. Farmacogenômica. Tamoxifeno

ABSTRACT

Breast cancer is the most incident cancer around the world, with high index of mortality, according to the markers expressed, it can be molecularly classified in luminal A or B (hormonal receptors), HER2 and basal-like breast cancer. Tamoxifen, the standard treatment of luminal breast tumors, is a prodrug that must be metabolized by liver enzymes, especially CYP2D6, to be therapeutically active. However, about 40% of patients do not respond to treatment with tamoxifen and there is evidence pointing to poor function of CYP2D6 causing underperformance of tamoxifen, although the data are conflicting. CYP2D6 is a highly polymorphic enzyme from P450 cytochrome, which is phenotypically classified according to enzyme activity in poor, intermediate, extensive and ultra metabolizers. The objective of this study was applying bioinformatics tools to assess the impact of missense mutations in CYP2D6 structure and function. The database CYPalleles was consulted to catalog the poor metabolizers alleles of CYP2D6 and their respective mutations. Thereafter, there were selected the alleles CYP2D6*7 and CYP2D6*14A, that contain only missense mutations. All polymorphisms in the selected poor metabolizers alleles were analyzed through seven algorithms to predict the structural and/or functional impact of each mutation over the protein. Subsequently, 3D models were constructed to evaluate the tertiary structure altered by the mutations. The results for the allele CYP2D6*7, showed that H324P mutation, which is the only one present in this allele, compromises the dynamics of the active site due to its location between two α -helices, resulting in high functional impact. The allele CYP2D6*14A has four mutations, three of them are present in various alleles but while R296C and S486T were not predicted as deleterious, P34S leads to a gain of one H-bond and severe protein stabilization, with function loss. The last polymorphism is G169R, that characterizes genotypically the allele and leads to a high structural and functional impact due to gain of solvent accessibility and severe enzyme destabilization, as it is present in an interface region between the monomeric chains. The *in silico* analyses showed that the deleterious effects of missense mutations in the structure and function of CYP2D6 may explain the resistance to tamoxifen treatment.

Key words: Bioinformatics. Breast cancer. CYP2D6. Pharmacogenomics. Tamoxifen

SUMÁRIO

1 INTRODUÇÃO.....	11
2 REVISÃO DE LITERATURA.....	12
2.1 Câncer de Mama.....	12
2.1.1 Tratamento para o Câncer de Mama.....	16
2.2 Citocromo P450 e Metabolização de Drogas.....	18
2.2.1 Farmacogenômica e CYP2D6.....	20
2.3 Bioinformática.....	24
3 OBJETIVOS.....	29
4 ARTIGO CIENTÍFICO.....	30
4.1 Material Suplementar.....	41
5 CONCLUSÕES E PERSPECTIVAS.....	43
REFERÊNCIAS.....	44
ANEXOS.....	50

1 INTRODUÇÃO

O câncer de mama é a neoplasia mais diagnosticada em mulheres, representa 25% de todos os tumores diagnosticados em 2012 e, com 1,67 milhões de novos casos, é o segundo câncer mais comum no mundo. Os cânceres de mama podem ser divididos em quatro subtipos moleculares: Luminal A, Luminal B, HER2+ e Basal-like. Este último, compreende um fenótipo específico denominado triplo negativo (TNBC - do inglês *triple negative breast cancer*) que não apresenta receptores hormonais expressos na superfície celular.

Aproximadamente 70% dos cânceres de mama expressam receptor de estrógeno (RE), mas de 30 a 50% das pacientes apresentam resistência a terapia endócrina com tamoxifeno, não respondendo ao tratamento (resistência intrínseca) ou recaindo após apresentar resposta inicial (resistência adquirida). Alguns mecanismos são propostos para justificar esta resistência: (i) alterações no receptor de estrógeno; (ii) alterações nos segundos mensageiros envolvidos na sinalização após ativação do RE e (iii) defeitos no metabolismo do tamoxifeno, que é um modulador seletivo dos RE e o padrão para tratamento endócrino. O tamoxifeno é uma pró-droga que após ser metabolizada no citocromo P450, principalmente pela CYP2D6, gera dois metabólitos ativos: endoxifeno e 4OH-tamoxifeno. O gene da CYP2D6 é altamente polimórfico, apresentando mais de 105 alelos descritos, que estão fenotipicamente classificados de acordo com a atividade metabólica em: pobres, intermediários, rápidos e ultra metabolizadores. Apesar de haver controvérsias, evidências indicam que os polimorfismos no gene da CYP2D6 podem ser a causa da má metabolização do tamoxifeno e, conseqüentemente, das falhas terapêuticas do medicamento.

O objetivo deste trabalho foi aplicar ferramentas de bioinformática para realizar análise *in silico* do impacto estrutural e funcional de mutações de sentido trocado (missenses) em alelos pobres metabolizadores da CYP2D6 e avaliar sua possível influência na resistência ao tratamento endócrino com tamoxifeno.

2 REVISÃO DE LITERATURA

O câncer é um problema de saúde pública, especialmente entre os países em desenvolvimento. A estimativa mundial realizada pelo projeto GLOBOCAN (IARC) em 2012 apontou que 60% dos 14 milhões de novos de cânceres e 70% dos 8 milhões de óbitos ocorram em países em desenvolvimento (FERLAY et al., 2012).

A estimativa do Instituto Nacional do Câncer (INCA) para o biênio 2016-2017 é de 600 mil novos casos de câncer. Excetuando-se o câncer de pele não melanoma, que corresponde a aproximadamente 180 mil casos novos, ocorrerão cerca de 420 mil casos novos de câncer. O perfil epidemiológico observado assemelha-se ao da América Latina e do Caribe, onde os cânceres de próstata (61 mil) em homens e mama (58 mil) em mulheres serão os mais frequentes (**Figura 1**). Para o ano de 2016 são esperados 57.960 novos casos de câncer de mama, com risco estimado em 56,20 casos a cada 100 mil mulheres. Nas regiões Sul, Sudeste, Centro-Oeste e Nordeste o câncer de mama deverá ser o primeiro mais frequente, sem considerar os tumores de pele não melanoma (MINISTÉRIO DA SAÚDE - BRASIL, 2015).

Figura 1: Distribuição proporcional dos dez tipos de câncer mais incidentes estimados para 2016 por sexo, exceto pele não melanoma

Localização Primária	Casos	%			Localização Primária	Casos	%
Próstata	61.200	28,6%		Homens	Mama feminina	57.960	28,1%
Traqueia, Brônquio e Pulmão	17.330	8,1%			Cólon e Reto	17.620	8,6%
Cólon e Reto	16.660	7,8%			Colo do útero	16.340	7,9%
Estômago	12.920	6,0%			Traqueia, Brônquio e Pulmão	10.890	5,3%
Cavidade Oral	11.140	5,2%			Estômago	7.600	3,7%
Esôfago	7.950	3,7%			Corpo do útero	6.950	3,4%
Bexiga	7.200	3,4%			Ovário	6.150	3,0%
Laringe	6.360	3,0%			Glândula Tireoide	5.870	2,9%
Leucemias	5.540	2,6%			Linfoma não Hodgkin	5.030	2,4%
Sistema Nervoso Central	5.440	2,5%			Sistema Nervoso Central	4.830	2,3%

*Números arredondados para múltiplos de 10.

Fonte: Instituto Nacional do Câncer (2015)

2.1 Câncer de Mama

O câncer de mama é o tipo de neoplasia mais comum em mulheres, correspondendo aproximadamente a 25% de todas as neoplasias em mulheres em todo o mundo. Podendo acontecer em qualquer uma das células da glândula mamária, o câncer de mama exibe propriedades morfológicas variáveis com diferentes perfis imunohistoquímicos e histopatológicos que apresentam prognóstico e curso clínico particulares (MAKKI, 2015).

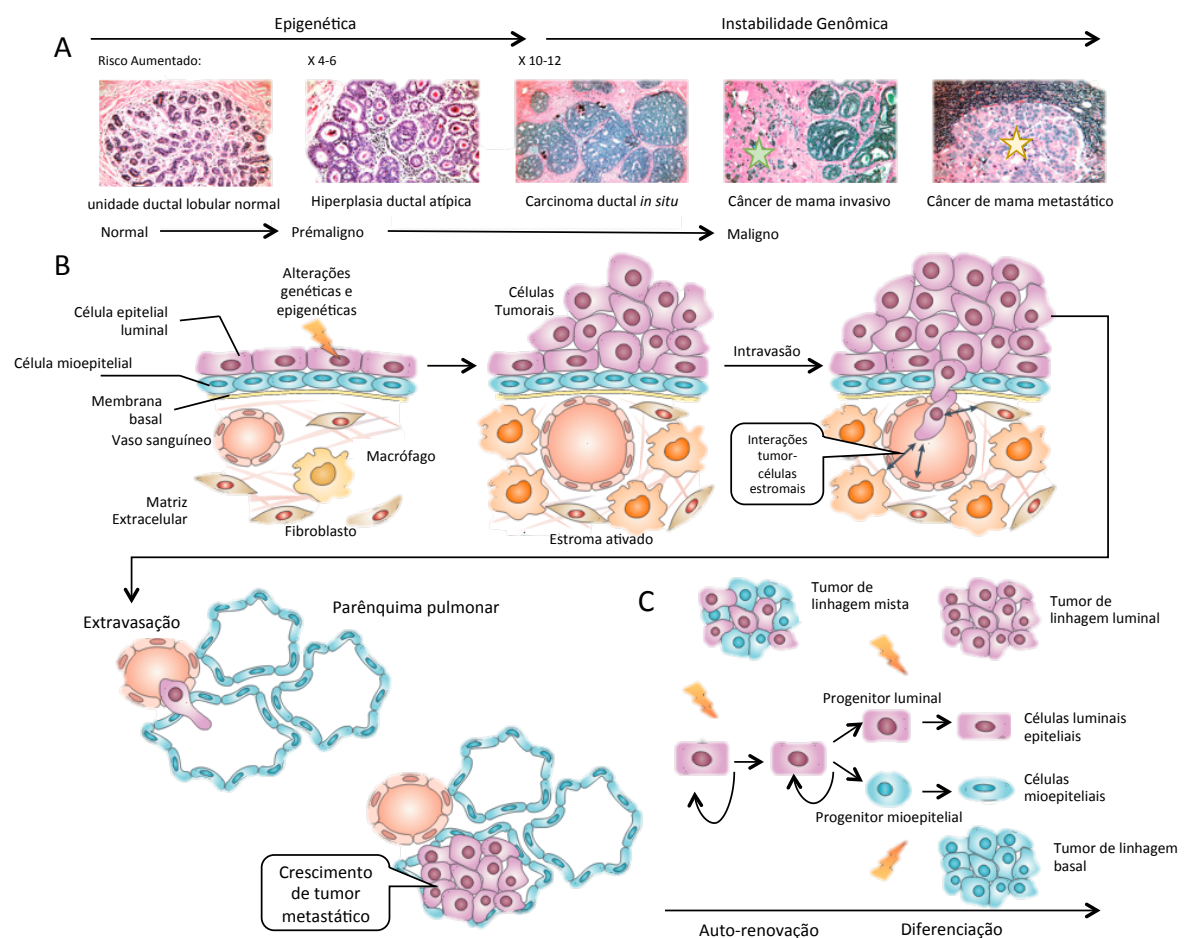
O câncer de mama é uma doença heterogênea geneticamente, que se desenvolve continuamente. O tecido mamário normal contém lóbulos e ductos que constituem um epitélio duplo de células luminais e estromais (**Figura 2A**). No tecido mamário normal as células epiteliais formam o ducto, estabelecem junções com as células vizinhas e com a membrana basal, constituindo uma forma de controlar a proliferação celular e garantir a manutenção da polaridade celular, através de inibição por contato. A hiperplasia ductal atípica é uma lesão pré-maligna que se caracteriza por ter camadas anormais de células epiteliais, sendo considerada uma lesão precursora do carcinoma ductal *in situ*, que não é invasivo. Neste tecido cancerígeno, as células ductais perdem a estrutura polarizada que é fornecida por uma matriz densa de colágeno, e desta forma, ficam propensas a proliferar e migrar (estrela verde). A medida que as células acumulam alterações epigenéticas e genéticas, o risco para o desenvolvimento de câncer de mama invasivo aumenta. Uma vez que o carcinoma se torne invasivo, o risco para desenvolvimento de metástase aumenta consideravelmente, sendo os linfonodos o sítio primário (VARGO-GOGOLA; ROSEN, 2007; VIDEIRA; REIS; BRITO, 2014).

Observa-se que no processo de transformação das células epiteliais da mama ocorre um acúmulo de alterações epigenéticas e genéticas, combinadas com interação anormal com o microambiente estromal. Este microambiente tumoral é complexo e envolve a matriz extracelular, células cancerígenas, além de diversos outros tipos de células, inclusive fibroblastos e células do sistema imune. Tamanha complexidade celular é responsável por garantir autossuficiência em fatores de crescimento, bem como manter o ambiente inflamatório e oxidante, favorecendo a angiogênese. A consequência é que células cancerígenas adquirem controle autônomo de sua proliferação, sobrevivência e migração. Dados estes fatores, as células cancerígenas são capazes de invadir a membrana basal e entrar nos vasos sanguíneos, através dos quais atingem diferentes tecidos e provocam metástases, como exemplificado no parênquima pulmonar (**Figura 2B**) (INSUA-RODRÍGUEZ; OSKARSSON, 2015; VARGO-GOGOLA; ROSEN, 2007; VIDEIRA; REIS; BRITO, 2014).

A hipótese de que apenas a minoria das células possuem potencial tumorigênico faz parte da teoria de que tumores contêm um conjunto de células especializadas com propriedades únicas (MAGEE; PISKOUNOVA; MORRISON, 2012) (**Figura 2C**). Estas células são chamadas células tronco cancerígenas pois têm habilidade de auto-renovação e diferenciação. É suposto que sejam responsáveis por gerar e manter o tumor estando associadas a progressão tumoral, quimiorresistência, recorrência e metástase (MAUGERI-

SACCÀ; DE MARIA, 2015; VARGO-GOGOLA; ROSEN, 2007; VIDEIRA; REIS; BRITO, 2014). Na **figura 2C** está esquematizado o papel da célula tronco mamária no desenvolvimento do câncer de mama.

Figura 2: Biologia do tecido mamário normal e cancerígeno. Em A observa-se a histologia do tecido mamário normal e sua transformação maligna onde a estrela amarela indica os linfonodos primários e a estrela verde a matriz colágena; em B evidencia-se a relação do tumor de mama com seu microambiente, levando a metástase; em C está esquematizado o papel da célula tronco mamária e o desenvolvimento do tumor.



Fonte: Adaptado de Vargo-Gogola & Rosen (2007)

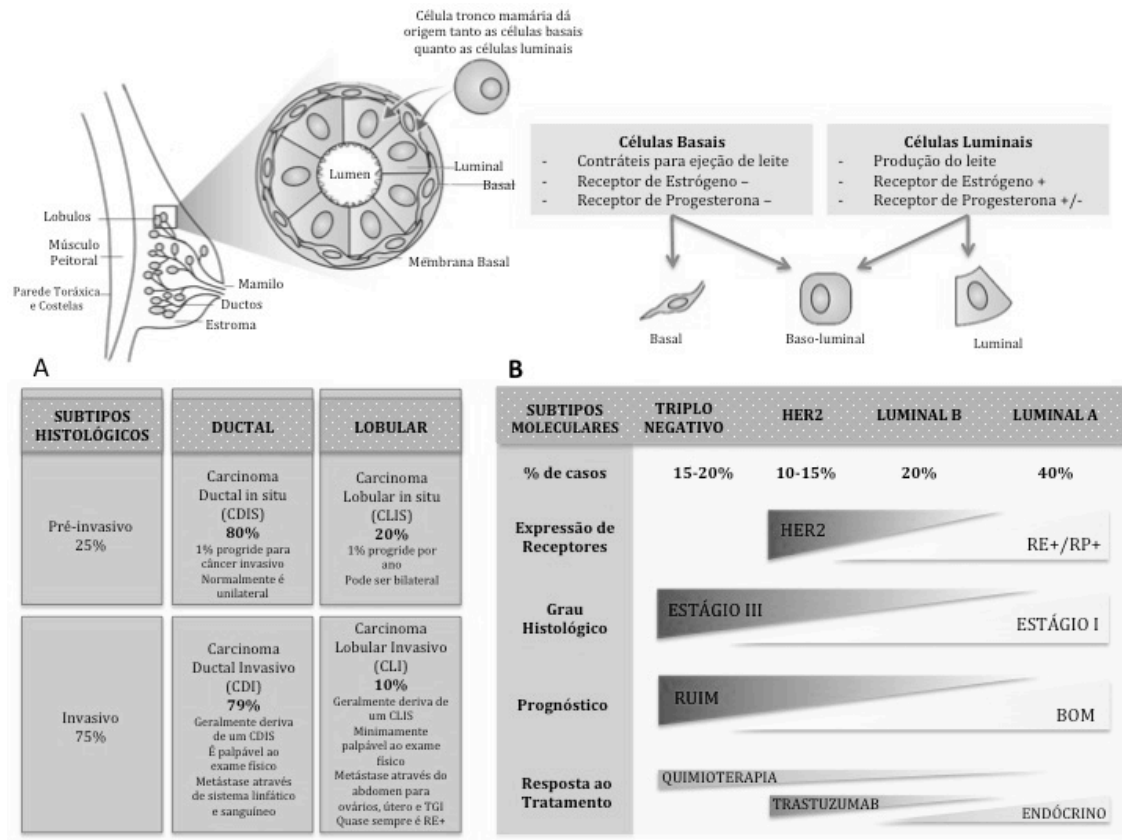
O câncer de mama é classificado em quatro estágios, considerando-se idade, comprometimento de linfonodos, grau e tamanho do tumor. Deste modo, o estágio I caracteriza uma massa tumoral pequena e localizada sem propriedades invasivas; nos estágios II e III o tumor é maior e apresenta fenótipo invasivo; os tumores avançados são classificados

como estágio IV. O comprometimento de linfonodos gera uma subclassificação em A, B e C (VIDEIRA; REIS; BRITO, 2014). Adicionalmente, 95% dos tumores mamários são adenocarcinomas (KUMAR; ABBAS; FAUSTO, 2010). Os cânceres originados dos ductos, são chamados ductais, enquanto aqueles que se originam nos lóbulos são chamados de carcinomas lobulares, ambos podem ser considerados *in situ* ou invasivos (**Figura 3A**) (MAKKI, 2015).

Perou fez análise de expressão gênica por microarranjo dos tumores de mama, classificando nos seguintes subgrupos: Luminal A e B, Basal-like, HER2 super expresso e mama normal (PEROU et al., 2000) (**Figura 3B**). O tipo basal-like é positivo para CK5/6, p-caderina, caveolinas 1 e 2, nestina, EGFR, entre outros; esse subtipo é frequentemente confundido com os tumores triplo negativos, no entanto, 30% dos tumores basais não tem o fenótipo triplo negativo (BADVE et al., 2011). O câncer de mama denominado de triplo negativo é caracterizado por não apresentar receptores de estrógeno (RE) e/ou progesterona (RP) e HER2. Por outro lado, os tumores Luminais expressam os RE e RP, mas o Luminal A não apresenta HER2 e tem baixo Ki67, um marcador de proliferação celular, em contraste ao tipo B que é HER2+ e tem alto Ki67 (LEE, 2016; PEROU et al., 2000). Outro subtipo foi posteriormente descrito, o *claudin-low* se refere a tumores que exibem propriedades de células tronco mamárias e mesenquimais, esse subtipo tem sido frequentemente associado a células resistentes à quimioterapia (O'REILLY et al., 2015). Esta classificação foi posteriormente comprovada por inúmeros estudos independentes e conduziu a mudança na conduta clínica, que passou a se basear neste novo parâmetro para tomar as decisões terapêuticas, bem como avaliar o prognóstico (EROLES et al., 2012).

A aplicação da tecnologia molecular no diagnóstico do câncer de mama possibilita direcionar mais eficientemente as decisões terapêuticas e estabelecer prognósticos mais precisos para as pacientes. Neste sentido, estão disponíveis no mercado alguns painéis moleculares, como o Oncotype DX® que avalia através de um score a resposta a terapia hormonal e a quimioterapia em pacientes que expressam receptor de estrógeno; e o MammaPrint que por sua vez, avalia o risco de recorrência e de metástases em pacientes recém diagnosticadas com câncer de mama (BERSE; LYNCH, 2015; MURAWA et al., 2014; STEIN et al., 2016)

Figura 3: Patogênese do câncer de mama e grau histológico VS subtipos moleculares.



Fonte: Adaptado de <http://www.pathophys.org/breast-cancer/>

2.1.1 Tratamento para o Câncer de Mama

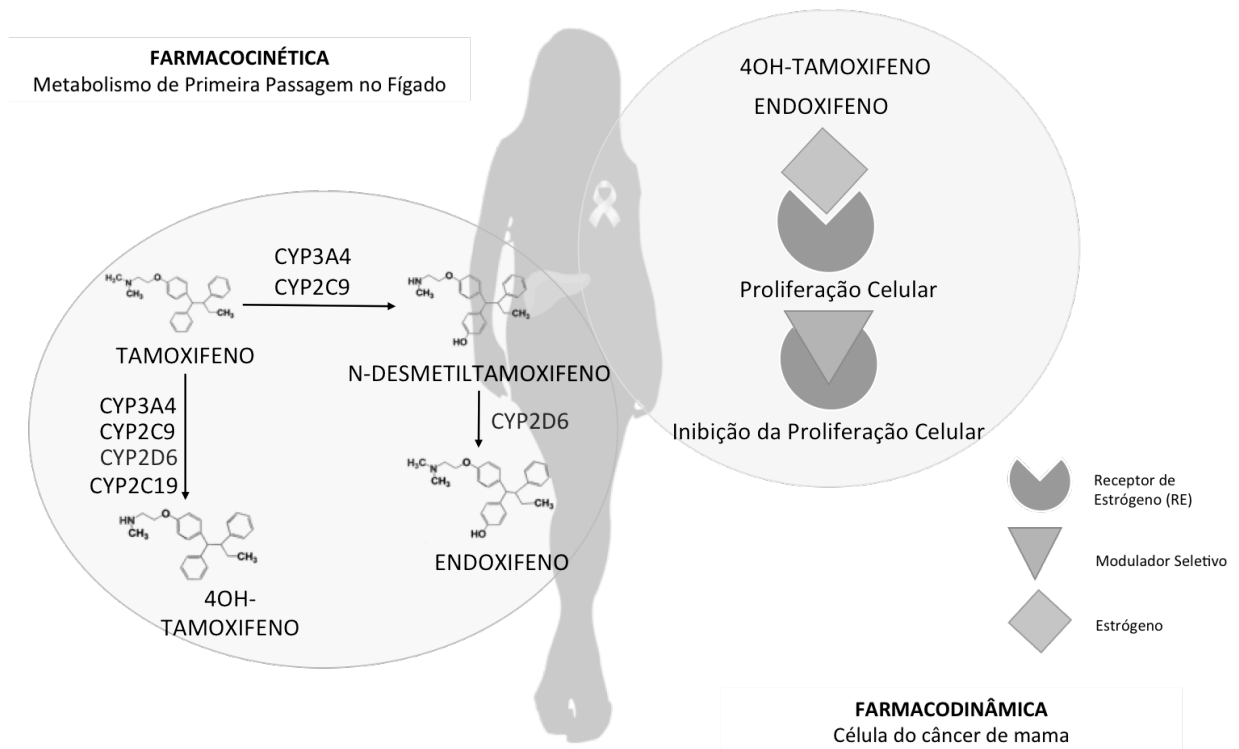
Avanços em métodos de rastreamento e diagnóstico precoce têm levado a redução significativa no número de recaídas e morte por câncer de mama. No entanto, tradicionalmente o tratamento de câncer de mama leva em consideração critérios clínicos e histopatológicos. Pode-se, portanto, pontuar três vias principais de tratamento: a quimioterapia, o tratamento endócrino e o tratamento com anticorpo monoclonal anti-HER2 (WESTBROOK; STEARNS, 2013). Aproximadamente 80% das pacientes de câncer de mama expressam receptores de estrógeno no tecido tumoral, isto é, se enquadram no grupo de câncer de mama tipo Luminal, sendo portanto direcionadas para o tratamento endócrino (LUMACHI; SANTEUFEMIA; BASSO, 2015).

Tratamento Endócrino

O tratamento endócrino do câncer de mama representa a terapia padrão para os tumores RE+ e atualmente se constitui de: (i) ablação ovariana; (ii) moduladores seletivos do receptor de estrógeno, que bloqueiam a ação do estrógeno no tumor e (iii) inibidores de aromatase, que inibem uma enzima chave na produção do estrógeno (LUMACHI; SANTEUFEMIA; BASSO, 2015). No entanto, de modo geral, o tamoxifeno é o medicamento mais utilizado como tratamento endócrino do câncer de mama, um modulador seletivo de receptores de estrógeno (SALADORES et al., 2014).

O tamoxifeno age como agonista parcial não-esteróide em alguns tecidos como fígado, útero e ossos, mas no cérebro e na mama age como um inibidor competitivo do receptor de estrógeno. Desta forma, o tamoxifeno bloqueia a sinalização desde o RE, inibindo a proliferação celular tumoral (LUMACHI; SANTEUFEMIA; BASSO, 2015) (**Figura 4**). No entanto, para exercer sua função terapêutica, o tamoxifeno precisa ser convertido a seus metabólitos ativos: o 4OH-tamoxifeno e endoxifeno, que têm cem vezes mais afinidade pelo receptor de estrógeno do que o tamoxifeno. Porém, endoxifeno é considerado o metabólito principal por ter concentração plasmática muito maior que a do 4OH-tamoxifeno (MÜRDTER et al., 2011).

De 30 a 50% das pacientes tratadas com tamoxifeno apresentam algum nível de resistência ao tratamento e não obtêm a resposta esperada (SALADORES et al., 2014). Existem diversos mecanismos que podem causar esta resistência ao tratamento endócrino, podendo-se classificar em três categorias principais: (i) alteração no alvo da droga, ou seja, o RE; (ii) alterações nos segundos mensageiros da via sinalizada pelo RE e (iii) defeitos de metabolismo (GROENENDIJK; BERNARDS, 2014). O metabolismo do tamoxifeno ocorre principalmente no fígado, através de enzimas que compõe o citocromo P450, originando seus metabólitos ativos após a catálise por diversas enzimas, especialmente CYP2D6 (SALADORES et al., 2014).

Figura 4: Farmacocinética e Farmacodinâmica do Tamoxifeno.

Fonte: A autora

2.2 Citocromo P450 e Metabolização de Drogas

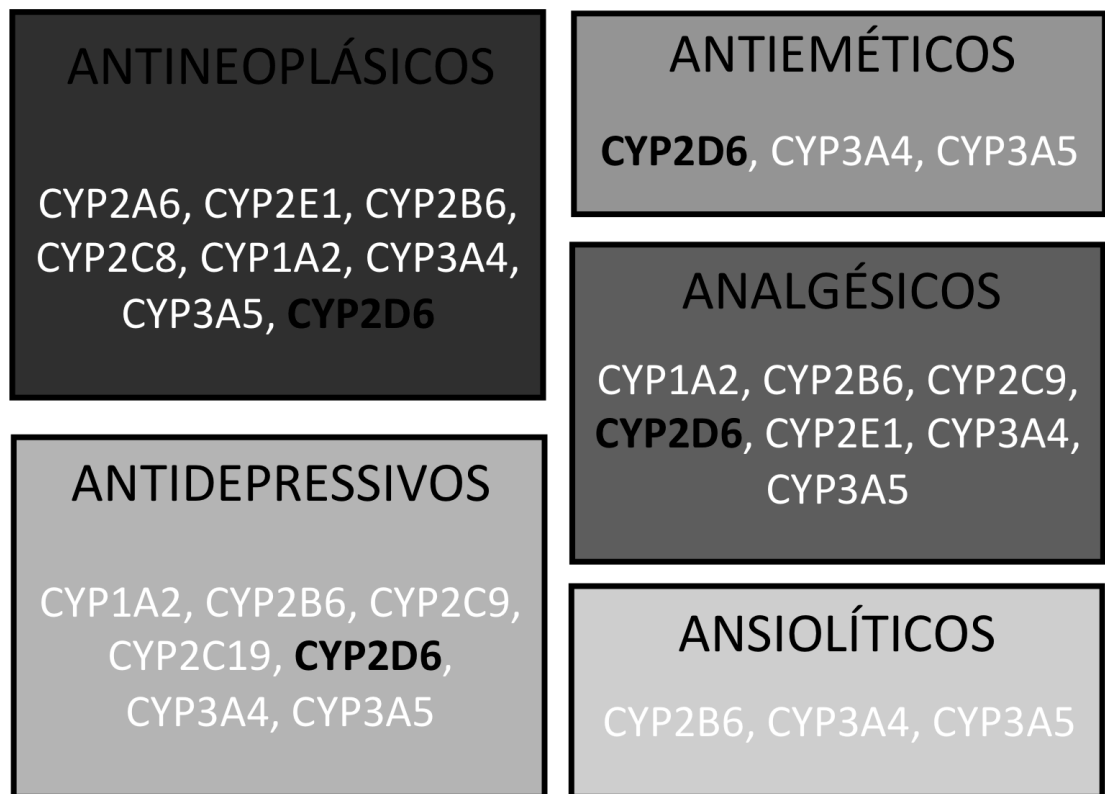
O citocromo P450 constitui uma família de enzimas capazes de catalisar a biotransformação oxidativa da maioria das drogas e de alguns xenobióticos lipofílicos, sendo considerado de grande relevância para farmacologia clínica. A maioria dos genes do P450 em humanos está agrupada de acordo com a similaridade das sequências genéticas em 18 famílias e 44 subfamílias, que desempenham funções endógenas específicas. No entanto, apenas as famílias de CYP1, 2 e 3 são responsáveis por metabolizar drogas e xenobióticos, e mesmo que haja sobreposição de especificidade por substratos entre as enzimas dessas famílias (**Figura 5**), várias drogas são metabolizadas a níveis clinicamente significantes por apenas uma ou poucas delas.

A nomenclatura das proteínas que compõe o P450 é complexa e adota regras específicas, descritas a seguir: primeiro o termo CYP designa enzimas do P450; o numeral arábico a seguir corresponde à família; é sucedido por uma letra maiúscula (subfamília) e o numeral arábico que vem por último, designando o peptídeo específico. Adicionalmente, a nomenclatura dos alelos de cada peptídeo é precedida por um asterisco, sendo estes nomeados

numerais arábicos seguidos de letra maiúscula (DALY et al., 1996; INGELMAN-SUNDBERG et al., 2000).

Existem diversos fatores que interferem na expressão e função destas CYPs, inclusive polimorfismos genéticos, e conhecer estes fatores é uma ferramenta importante no entendimento da farmacogenética e resposta terapêutica das drogas (ZANGER; SCHWAB, 2013).

Figura 5: Enzimas das famílias 1, 2 e 3 do citocromo P450 e algumas classes de drogas metabolizadas.



Fonte: Figura da autora, com informações disponíveis em Zanger & Schwab (2013)

O gene mais polimórfico entre os das CYPs responsáveis por metabolizar compostos exógenos é o da CYP2D6, para o qual existem mais de 100 variantes alélicas descritas (HERTZ et al., 2015; INGELMAN-SUNDBERG; SIM, 2010). A CYP2D6 é responsável por metabolizar cerca de 25% de todas as drogas exógenas, sendo a enzima mais importante no metabolismo do tamoxifeno (HERTZ et al., 2015).

2.2.1 Farmacogenômica e CYP2D6

A farmacogenômica tem por objetivo identificar variantes genômicas que influenciam o efeito das drogas e proporcionar a aplicação clínica da medicina personalizada (RELLING; EVANS, 2015). Informações do projeto 1000Genomas mostram que cada indivíduo carrega aproximadamente 300 variantes genéticas associadas a perda total ou parcial da função dos produtos gênicos. O fato de um único gene estar associado a várias doenças enquanto uma única doença pode ser causada por mutações em vários genes faz com que a relação entre o genótipo-fenótipo seja muito mais complexa do que se poderia prever. Adicionalmente, as mutações não têm efeitos idênticos em diferentes indivíduos, pois fatores individuais como interações gênicas, proteicas e metabólicas, bem com fatores ambientais exercem grande influência nos fenótipos observados (ZHANG; KUIVENHOVEN; GROEN, 2015).

É possível, contudo, observar que a farmacogenética do indivíduo afeta de duas formas principais o metabolismo e ação dos medicamentos. A primeira é através de alterações na farmacocinética – que consistem na absorção, distribuição, metabolização e eliminação da droga; e a segunda através de modulação da farmacodinâmica da droga - modificando o alvo da droga ou alterando as vias biológicas que alteram a sensibilidade aos efeitos farmacológicos da droga (RELLING; EVANS, 2015).

Na farmacogenômica do câncer de mama os polimorfismos da CYP2D6 têm sido considerados como biomarcadores emergentes na predição de sucesso ao tratamento endócrino (KALIA, 2015). A CYP2D6 é relativamente pouco expressa quando comparada a outras enzimas do citocromo P450, no entanto metaboliza até 25% das drogas utilizadas na clínica. Estão inclusos em seus substratos virtualmente todas as classes terapêuticas tais como: antiarrítmicos, antidepressivos, antipsicóticos, beta-bloqueadores, analgésicos e antineoplásicos (BAUMANN, 2015). Neste contexto é importante considerar o fato da polifarmácia, ou polimedicamento – administração de vários medicamentos simultaneamente, um proeminente fator ambiental que interfere na atividade da CYP2D6. O fenômeno denominado de fenoc conversão é caracterizado pela perda de metabolização ainda que não existam polimorfismos genéticos relacionados. Usualmente a fenoc conversão da CYP2D6 ocorre por causa da concomitância na administração de substratos e inibidores (LISBETH et al., 2016) (**Figura 6**).

Figura 6: Principais substratos e inibidores da CYP2D6

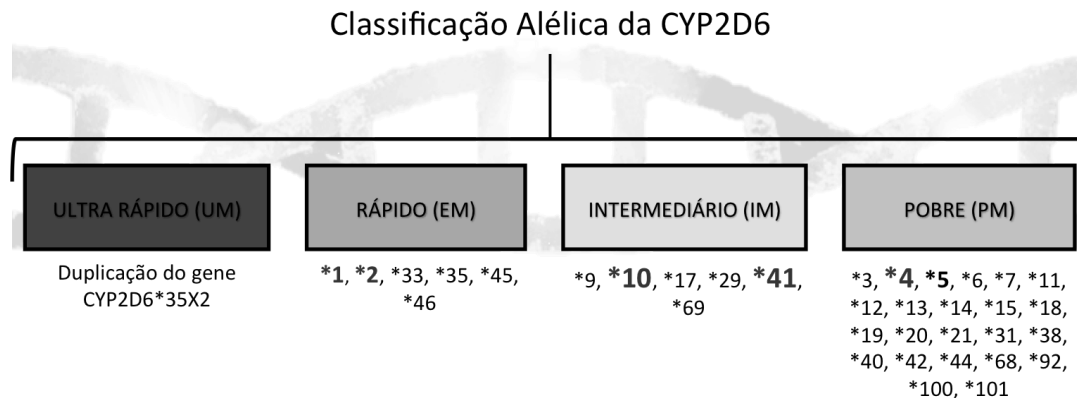
Antidepressivos	Antipertensivos	Antipsicóticos	Antieméticos	Analgésicos	Inibidores
☐ Amitriptilina	☐ Debrisoquina	☐ Aripiprazol	☐ Dolasetrona	☐ Codeína	☐ Bupropiona
☐ Atomoxetina	☐ Metoprolol	☐ Risperidona	☐ Ondansetrona	☐ Dihidrocodeína	☐ Celecoxib
☐ Desipramina	☐ Propranolol	☐ Tioridazina	☐ Prometazina	☐ Hidrocodona	☐ Haloperidol
☐ Duloxetina	☐ Timolol	☐ Zuclopentixol	☐ Tropizetrona	☐ Tramadol	☐ Metadona
☐ Imipramina	☐ Carvedilol	☐ Clorpromazina			☐ Paroxetina
☐ Mianserina					☐ Quinidina
☐ Mirtazapina					☐ Fluoxetina
☐ Nortriptilina					☐ Flecainida
☐ Paroxetina					
☐ Venlafaxina					

Fonte: A autora

Para avaliar a associação entre a perda da função da CYP2D6 por polimorfismos genéticos e a resposta ao tratamento endócrino do câncer de mama, inúmeros estudos tem sido desenvolvidos, sem no entanto, encontrar consenso (GOETZ et al., 2013; RAE et al., 2012; REGAN et al., 2012; SCHROTH et al., 2014).

O gene CYP2D6 está localizado no braço longo do cromossomo 22, é altamente polimórfico com mais de 105 variantes alélicas descritas, sendo aproximadamente 40 SNPs (*Single Nucleotide Polymorphisms*, do inglês) (KIM et al., 2013). Os diferentes genótipos da CYP2D6 refletem em uma caracterização fenotípica dos indivíduos de acordo com a atividade da CYP2D6: (i) pobres metabolizadores – PM, que tem dois alelos não funcionais; (ii) metabolizadores intermediários – IM, um alelo funcional ou dois alelos de atividade reduzida; (iii) metabolizadores extensivos – EM, que tem dois alelos funcionais – considerados normais; e (iv) metabolizadores ultrarrápidos – UM, que tem duplicação de alelos funcionais (**Figura 7**).

Figura 7: Classificação alélica da CYP2D6 e distribuição dos alelos representados por *N (onde N é o respectivo alelo). Os alelos mais frequentes estão em destaque .



Fonte: A autora, com informações da base de dados CYPalleles

Os efeitos farmacológicos decorrentes dos polimorfismos da CYP2D6 variam de acordo com o substrato, mas a consideração mais importante nesse sentido é se a droga é inativada ou ativada pela oxidação catalisada pela CYP2D6. No caso de drogas que são inativadas, em indivíduos PM é esperado que a droga se acumule resultando em maior risco de efeitos adversos e toxicidade. Em contraste os indivíduos UM deverão perder parte do efeito da droga, dado o rápido *clearance* da droga ativa. Para as pró-drogas ativadas pela CYP2D6 a situação é consequentemente inversa, os PM experimentam falha terapêutica e os UM têm mais efeitos colaterais (BAUMANN, 2015).

Para o tamoxifeno foi observada a existência um efeito gene-dose dependente, onde indivíduos que tem fenótipo PM ou IM para CYP2D6 obtêm concentração plasmática de endoxifeno até 75% menor do que aqueles com fenótipo metabolizador extensivo (BINKHORST et al., 2015). Em populações caucasianas a mutação funcionalmente relevante mais frequente é a 1846G>A, do alelo pobre metabolizador CYP2D6*4. Este polimorfismo resulta em um RNA mensageiro fruto de *splicing* alternativo com uma base a mais, levando a formação de código de parada prematuro e como consequência, uma proteína truncada e de atividade nula. Na população asiática o alelo CYP2D6*10 é o mais frequente ocorrendo em 50% da população. Indivíduos homozigotos para o alelo *10 ou indivíduos heterozigotos com um alelo nulo tem atividade *in vivo* substancialmente reduzida se comparados aos EM, sendo classificados como IM. Outro alelo reportado majoritariamente para uma população em específico foi o *17, em 30% dos africanos. O CYP2D6*17 possui um haplótipo complexo com diversas substituições de aminoácido. Dentre estas, a troca de treonina por isoleucina na

posição 107 foi relacionada com atividade reduzida e alterações na preferência por substratos *in vitro*. (BAUMANN, 2015)

Um estudo com objetivo de analisar as variabilidades interétnicas dos alelos da CYP2D6 mostrou que na população americana, há maior frequência do alelo *4. Este mesmo estudo aponta que a população brasileira é genotipicamente similar a de Cuba e dos Estados Unidos (LLERENA et al., 2014). Dois estudos realizados na população brasileira mostram contraste entre as regiões nordeste e sul, no qual do Sul do Brasil (Porto Alegre) o alelo *4 teve frequência de 18% e o *10 de 1% (ANTUNES et al., 2012); no Nordeste (São Luís) o alelo *4 teve frequência de 25% e o alelo *10 de 14% (MARTINS et al., 2014).

Em uma tentativa de racionalizar o uso de medicamentos algumas tecnologias de genotipagem da CYP2D6 já foram desenvolvidos, o mais proeminente dentre eles é o Amplichip® (ROCHE Molecular Diagnostics), aprovado pela FDA (*Food and Drugs Administration* – EUA) em 2005. A tecnologia é baseada em microarranjo e emprega cinco etapas: (i) amplificação por reação em cadeia da polimerase (PCR) do DNA purificado; (ii) fragmentação e marcação do produto amplificado; (iii) hibridização dos fragmentos com sondas do microarranjo; (iv) leitura do microarranjo e (v) determinação do genótipo e predição do fenótipo da CYP2D6 e CYP2C19 (LEON; SUSCE; MURRAY-CARMICHAEL, 2006). O Amplichip é capaz de genotipar 33 alelos da CYP2D6, mas seu custo permanece elevado, dificultando sua implementação no diagnóstico clínico.

Diversos estudos multicêntricos com número considerável de pacientes já foram realizados, mas sem encontrar consenso. Possivelmente a causa para as discordâncias está relacionada a não conformidade em diversos critérios importantes, pode-se pontuar: número de variantes alélicas consideradas, fonte de DNA para testes de genotipagem, o tratamento quimioterápico feito em paralelo, além da duração do tratamento com tamoxifeno (BINKHORST et al., 2015).

Considerando-se a importância da CYP2D6 no metabolismo de drogas exógenas e a falta de concordância nos estudos *in vitro*, torna-se importante utilizar outras abordagens *in silico* para compreender o papel dos polimorfismos da CYP2D6 na falha ao tratamento com tamoxifeno.

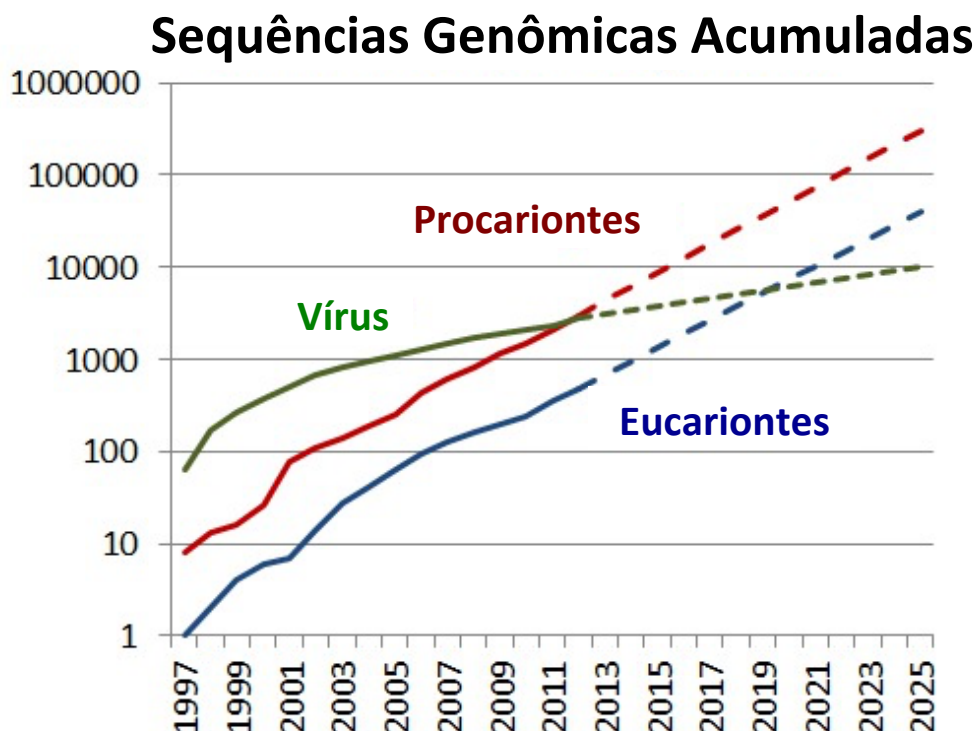
A bioinformática, através de recursos computacionais aplicados aos dados biológicos, pode contribuir com a farmacogenômica aplicando ferramentas e abordagens que vão desde a

mineração de dados depositados em bases de dados públicas, o desenvolvimento e aplicação de algoritmos de predição até análises 3D de interação proteína-proteína e proteína-droga (CHANG, 2015; SEAL et al., 2014; VAN NESTE; VAN CRIEKINGE, 2015; VERLI, 2014).

2.3 Bioinformática

Na década de 1990 foi lançado o Projeto Genoma Humano. Aproximadamente 14 anos depois a sequência completa do genoma humano com mais de 3 bilhões de pares de bases estava disponível. Desde então, os dados genômicos têm sido coletados em proporções cada vez maiores (CAPRIOTTI et al., 2012). Numa relação inversamente proporcional, com a queda dos custos de sequenciamento, houve um aumento de sequências e polimorfismos depositados nas bases de dados (**Figura 8**).

Figura 8: Análise retrospectiva e prospectiva do número de genomas completos sequenciados.



Fonte: <http://sulab.org/2013/06/sequenced-genomes-per-year/>

O projeto genoma já reportou mais de 84.000.000 de SNPs, 3.000.000 de pequenas inserções ou deleções e 60.000 variantes estruturais em 2.504 indivíduos de 26 populações diferentes. Adicionalmente, foi reportado que cada indivíduo carrega aproximadamente 300 variantes que levam a produtos gênicos com perda total ou parcial de função (ZHANG; KUIVENHOVEN; GROEN, 2015).

As variantes genéticas podem ser classificadas como SNP quando ocorrem em pelo menos 1% da população (FRANK, 1999). Os SNPs podem ocorrer em regiões codificantes de genes, em regiões não codificantes ou em regiões intergênicas. Os SNPs que se encontram em regiões não codificantes podem ter consequências no *splicing*, na ligação de fatores de transcrição ou na sequência de RNAs não codificantes, que exercem funções regulatórias. As mutações sinônimas ou silenciosas geram a mesma sequência polipeptídica. Os SNPs não sinônimos podem ser missense ou de perda de sentido, onde a de perda de sentido leva a um código de parada prematuro e a missense a uma substituição de aminoácido (WU et al., 2015).

As mutações de sentido trocado podem alterar a função da proteína de diversas formas: (i) mutações que ocorrem em resíduos funcionais - como os do sítio ativo ou envolvidos em interação proteína-proteína, podendo causar ganho ou perda de função e consequentemente alteração na via molecular envolvida; (ii) as que podem alterar a estabilidade da proteína seja desestabilizando (aumento da proporção de proteína enovelada) ou estabilizando (diminuição da proporção de proteína enovelada), o que também deverá ocasionar perda funcional; (iii) mutações que podem causar agregação da proteína final, impedindo sua correta função; e (iv) as que podem levar a alterações nos processos pós-traducionais, como perda/ganho de sítios de glicosilação (WORTH et al., 2007)

Diante desta perspectiva, sabe-se que os SNPs têm grande influência em como os humanos respondem a diferentes situações que vão desde a exposição a patógenos e agentes químicos até aos medicamentos e vacinas. Desta forma, o estudo destes polimorfismos tem grande importância na pesquisa biomédica e no desenvolvimento de medicamentos (CAPRIOTTI et al., 2012; WANG et al., 2012; WU et al., 2015).

Considerando-se que substituições de aminoácidos podem levar a alterações críticas na estrutura da proteína enovelada, instabilidade estrutural e até perda funcional, diversos métodos têm sido desenvolvidos para prever os efeitos das trocas de aminoácidos. Pode-se destacar duas classes, não necessariamente independentes: (i) predição de efeito funcional e (ii) predição de impacto na estabilidade proteica após mutações de sentido trocado

(CAPRIOTTI et al., 2012). De modo geral, estes métodos se baseiam em algoritmos, que são por definição, uma sequência lógica de instruções necessárias para executar determinada tarefa (VERLI, 2014). Os algoritmos aplicam métodos de classificação como *Support Vector Machine* (SVM), *Neural Network* (NN) ou *Random Forest* (RF) para definir se um SNP tem provável efeito deletério nas proteínas (KANDOI; ACENCIO; LEMKE, 2015; KUMAR et al., 2014).

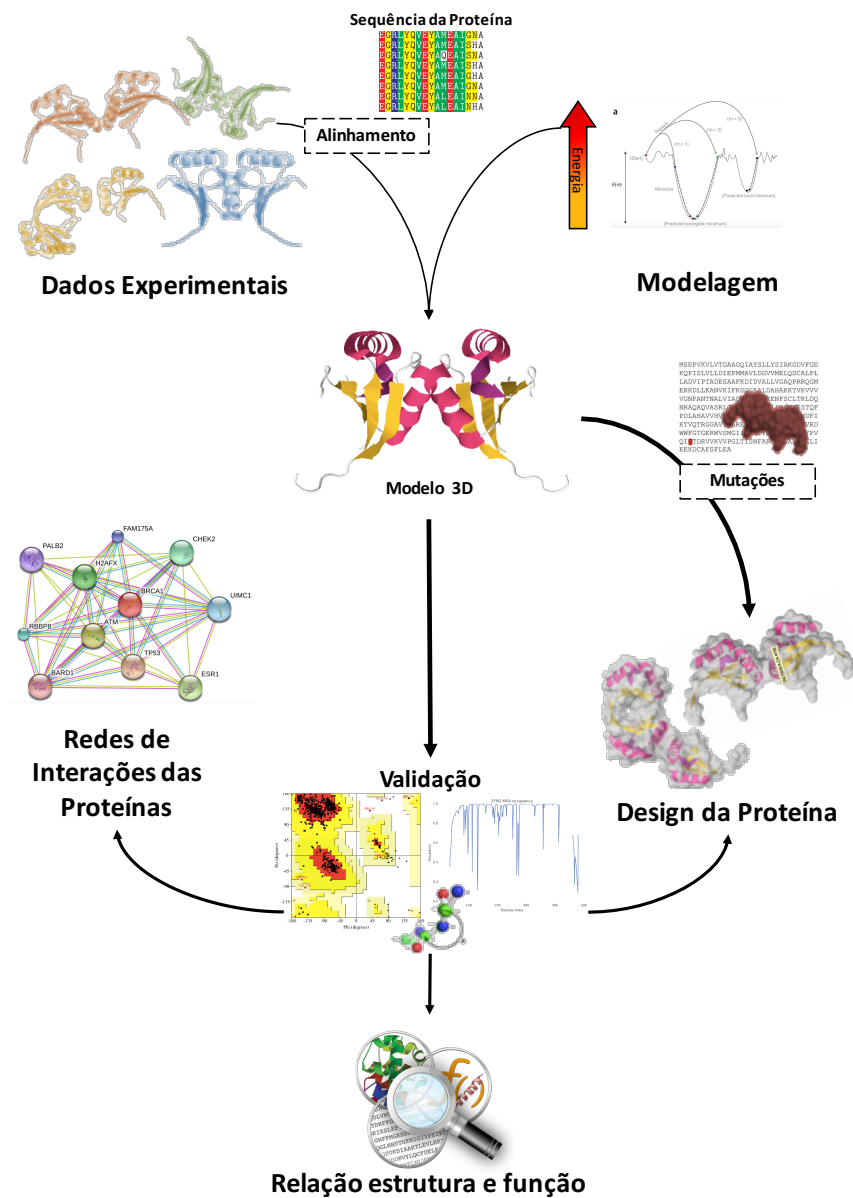
Dentre estes métodos, o mais clássico e consequentemente mais utilizado é o SVM, que foi desenvolvido numa base de aprendizagem estatística teórica para maximizar as margens de separação entre exemplos de duas classes projetadas em um hiperespaço (LI et al., 2009). Trata-se de um método de reconhecimento de padrões do tipo *kernel-based* que utiliza a aprendizagem de máquinas (*machine learning*). A diferença do SVM para os demais métodos de classificação é que os modelos de SVM dependem apenas das amostras próximas dos limites entre duas (ou mais) classes – ou seja, as amostras nas margens de cada classe, que são chamadas de vetores de suporte (*support vectors*) (LIU et al., 2013).

Os demais métodos utilizam todas as amostras do conjunto de dados para determinar os limites entre as classes (GUNN, 1998; LIU et al., 2013). O RF é um conjunto de árvores em um espaço vetorial de múltiplas dimensões de variáveis de um objeto classificado, seu resultado é baseado em uma média dos resultados independentes de cada árvore (LI et al., 2009; LIU et al., 2013). O método de *Neural Network*, ou Redes Neurais, mimetiza o funcionamento do cérebro humano e tem sido cada vez difundido por dois motivos principais: (i) como as propriedades das sequências proteicas estão distribuídas em hiperespaços com características complexas, normalmente é difícil encontrar modelos satisfatórios usando abordagens estatísticas ou parametrizadas; (ii) métodos baseados em NN são capazes de processar valores contínuos apresentados ao modelo (CAO; XIONG, 2014).

Apesar de utilizarem métodos diversos para classificar os polimorfismos, os algoritmos partem de dados biológicos similares para realizar as análises. Alguns baseiam suas análises exclusivamente em dados de conservação obtidos através de alinhamento de múltiplas sequências, como o Mutation Assessor e o SIFT. Outras abordagens, como as aplicadas pelo PolyPhen2, SNPs&GO e MutPred, combinam informações de homologia com vários tipos de anotações funcionais e estruturais, tais como: (i) propriedades físico-químicas dos aminoácidos envolvidos na troca; (ii) localização de regiões funcionais; (iii) estrutura secundária e (iv) topologia da proteína (FROUSIOS et al., 2013).

A utilização em conjunto com os algoritmos de predição e de análises estruturais 3D amplia as margens de predição para uma compreensão tridimensional dos efeitos dos polimorfismos missenses. A bioinformática estrutural (**Figura 9**) foca na análise de predição da estrutura 3D de macromoléculas biológicas e na identificação de relações entre a estrutura e a função, prevendo as interações moleculares que fazem parte de inúmeras vias de um sistema biológico (CHOONG; TYE; LIM, 2013; FISER, 2010; LI; WEI, 2015).

Figura 9: Aplicações da bioinformática estrutural. Exemplo de *workflow* a partir de análises experimentais de proteômica.



Fonte: A autora

A utilização de algoritmos de predição e *softwares* de modelagem estrutural para aplicações farmacogenéticas tem potencial de criar ferramentas personalizadas para o diagnóstico, prognóstico e tratamento de doenças, inclusive do câncer. Diversos estudos têm sido realizados utilizando ferramentas de bioinformática para fins farmacogenéticos (ALENCAR; LOPES, 2010; GEORGE PRIYA DOSS et al., 2013; SHANTHI; RAJASEKARAN; RAMANATHAN, 2014), inclusive um estudo que focou na análise do citocromo P450. Para este estudo os genes codificantes de mais de 40 enzimas que compõe o P450 foram submetidos a análise *in silico* para avaliar as principais variantes deletérias de cada gene. Foram analisadas 18 famílias de CYPs através dos algoritmos SIFT e PolyPhen. Dos 791 polimorfismos analisados 38 foram preditos como deletérios pelo SIFT, enquanto para o PolyPhen que 338 seriam deletérios. Neste estudo, o gene CYP2D6 apareceu entre os 10 com mais polimorfismos deletérios, tendo as variantes G42R, G169R, S311L, H324P, R343G, Y355C e R365H sido preditas como as mais deletérias tanto pelo SIFT como pelo PolyPhen. (WANG; LI; ZHOU, 2009)

Apesar de ter sido um estudo pioneiro na área, este trabalho com enzimas do P450 não avaliou a relação entre o impacto dos polimorfismos encontrados e a caracterização fenotípica dos alelos. Dada a importância da CYP2D6 para o metabolismo do tamoxifeno e considerando-se que não há concordância sobre o impacto de seus polimorfismos com as falhas na terapia com tamoxifeno, fica evidente a necessidade de utilizar novas técnicas e abordagens para responder estas questões.

3 OBJETIVOS

3.1 Objetivo Geral

- Analisar o impacto de mutações missenses na estrutura e função de alelos pobres metabolizadores da CYP2D6

3.2 Objetivos Específicos

- Catalogar as mutações presentes em alelos pobres metabolizadores da CYP2D6
- Selecionar alelos pobres metabolizadores contendo apenas mutações missenses
- Validar as mutações missenses em diferentes bases de dados
- Estabelecer os algoritmos de predição estrutural e funcional a serem utilizados
- Determinar o sistema binário de classificação das mutações
- Avaliar a estrutura 3D alterada pelas mutações não sinônimas

4 ARTIGO CIENTÍFICO

Pharmacogenomics



[Aims & Scope](#)
[Editorial Board Members](#)
[Indexing](#)

Impact factor: 3.218 (2014)

Current Issue

[View Volume 17, Issue 4 \(2016\)](#)

Purchase:

[View pricing](#)

[RSS](#) [What is RSS?](#)

Periódicos Qualis

Dados para Consulta

*Evento de Classificação:

Qualis 2014

Área de Avaliação

☒ CIÊNCIAS BIOLÓGICAS I

ISSN:

☐

Título:

☒ Pharmacogenomics

Classificação:

☒ -- SELECIONE --

Consultar

Cancelar

Periódicos

ISSN	Título	Área de Avaliação	Classificação
1462-2416	Pharmacogenomics (London)	CIÊNCIAS BIOLÓGICAS I	A2

Short Communication

For reprint orders, please contact: reprints@futuremedicine.com

Pharmacogenomics



Evaluating the impact of missenses mutations in *CYP2D6**7 and *CYP2D6**14A: does it compromise tamoxifen metabolism?

CYP2D6 is a high polymorphic enzyme from P450, responsible for metabolizing almost 25% of drugs. The distribution of different mutations among *CYP2D6* alleles has been associated with poor, intermediate, extensive and ultra-metabolizers. **Aim:** To evaluate how missenses mutations in *CYP2D6**7 and *CYP2D6**14A poor metabolizer alleles affect *CYP2D6* stability and function. **Materials & methods:** CYPalleles database was used to collect polymorphisms data present in 105 alleles. We selected only poor metabolizers alleles that presented exclusively missenses mutations. They were analyzed through seven algorithms to predict the impact on *CYP2D6* structure and function. **Results:** H324P, the unique mutation in *CYP2D6**7, has high impact in enzyme function due to its occurrence between two alpha-helices involved in active site dynamics. G169R, a mutation that occurs only in *CYP2D6**14A, leads to the gain of solvent accessibility and severe protein destabilization. **Conclusion:** Our *in silico* analysis showed that missenses mutations in *CYP2D6**7 and *CYP2D6**14A cause *CYP2D6* dysfunction.

First draft submitted: 4 December 2015; Accepted for publication: 28 January 2016; Published online: 4 April 2016

Keywords: algorithms • bioinformatics • *CYP2D6* • *CYP2D6**7 • *CYP2D6**14A • poor metabolizers • tamoxifen

Breast cancer is the second most common cancer in the world and, by far, the most frequent cancer among women with an estimated 1.67 million new cancer cases diagnosed in 2012 (25% of all cancers). Breast cancer ranks as the fifth cause of death from cancer overall (522,000 deaths), while it is the second most frequent cause of cancer death in women from America's region (92,058 deaths, 14.9% of total) [1].

Tamoxifen, is the standard treatment for estrogen receptor (ER)-positive breast cancer (70–80% of the cases), it is a compound known as selective ER modulator that acts as ER antagonist in breast tissue [2]. However, endocrine therapy using tamoxifen is known to fail in up to 50% of patients, due to alterations in estrogen receptor; alterations in pathways signaled by estrogen receptor; and

bypass mechanisms, that includes metabolic defects, mainly by *CYP2D6* dysfunction [3].

Tamoxifen has a weak affinity to the ER and is regarded as a prodrug, extensively metabolized by several members of the CYP450 family and Phase II conjugation enzymes. Multiple CYP enzymes are involved in the tamoxifen metabolism, including CYP3A4, CYP2B6, CYP2C9, CYP2C19 and *CYP2D6*. Endoxifen and 4-hydroxytamoxifen, which are the main metabolites, have equivalent anti-estrogenic potentials and are 30–100 times more active than tamoxifen. Because of the five- to tenfold higher plasma concentrations of endoxifen compared with 4-hydroxytamoxifen, endoxifen is believed to be the principal active metabolite [4]. It is also the only metabolite that binds to ER and leads to proteasomal degradation, decreasing ER levels [5].

Maria ACSM Borba^{*1}, Renato P Melo-Neto¹, Glauber M Leitão^{1,2}, Carlos HM Castelletti^{1,3}, José L Lima-Filho^{4,5} & Danyelly BG Martins^{1,4}

¹Molecular Prospection and Bioinformatics Group (ProspecMol) – Laboratory of Immunopathology Keizo Asami (LIKA), Federal University of Pernambuco (UFPE), Av. Prof. Moraes Rego 1235, 50670-901, Cidade Universitária, Recife, PE, Brazil

²Clinical Hospital – Federal University of Pernambuco (UFPE), Av. Prof. Moraes Rego 1235, 50670-901, Cidade Universitária, Recife, PE, Brazil

³Agronomic Institute of Pernambuco (IPA), Av. General San Martin 1371, 50761-000, Bongi, Recife, PE, Brazil

⁴Biochemistry Department, Federal University of Pernambuco (UFPE), Av. Prof. Moraes Rego 1235, 50670-901, Cidade Universitária, Recife, PE, Brazil

⁵Laboratory of Immunopathology Keizo Asami (LIKA), Federal University of Pernambuco (UFPE), Av. Prof. Moraes Rego 1235, 50670-901, Cidade Universitária, Recife, PE, Brazil

*Author for correspondence: mborba@prospecmol.org

Future
Medicine  part of 

CYP2 family contains 16 full-length genes, spread over different chromosomes and organized in multi-gene clusters, which all have nine exons and eight introns. CYP genes involved in pharmacological response are highly polymorphic, notably CYP2D6 [6], a tetramer heme-containing structure that acts over substrates that have a basic nitrogen and a planar aromatic ring [7]. It plays an important role in the metabolism of almost 25% of drugs, including analgesics, antidepressants and antineoplastic [6,8]. According to CYP Allele Nomenclature Database [9], there are 105 alleles for CYP2D6. These alleles have been associated with four drug metabolism phenotypes: poor, intermediate, extensive and ultra-metabolizers [10]. However, this association is still unclear to be applied in the guidelines for breast cancer treatment [11]. Some multicenter studies revealed a clear association between CYP2D6 alleles and the success of tamoxifen treatment [12,13], while many others demonstrated the failure of tamoxifen regardless of CYP2D6 genotype [14,15].

Nowadays, there is an intuitive rationale for using pharmacogenomics to guide medical treatment. So, the understanding of CYP2D6 alleles value in predicting the efficiency of tamoxifen through pharmacogenomics markers could improve the implementation of personalized medicine in the clinical oncology [16]. However, there is no well-established correlation between phenotypic and clinical data, despite the large volume of genomic data.

Therefore, bioinformatics approaches have been used as powerful tools to handle and understand genomic data [17,18]. Multiple algorithms were developed for predicting the structural, functional and evolutionary impact of genomic mutations based on protein sequence and 3D structure data. These tools provide support for experimental validation of disease-related alleles, emphasizing changes that could affect the structure and function of the protein [19,20]. There is also a class of methods, referred as template-based modeling, which relies on detectable similarity spanning most of the modeled sequence and at least one known structure [21]. As a small change in protein sequence usually results in a change in its 3D structure, the integration of algorithms data and 3D models leads to a more reliable conclusion.

The aim of this study was to evaluate the impact of missense mutations in CYP2D6*7 and CYP2D6*14A poor metabolizer alleles, using *in silico* approaches for predicting the structural and functional impact on the CYP2D6 molecule.

Materials & methods

Dataset construction

CYPalleles database [9] was used to collect polymorphisms data present in 105 alleles. Only 33 alleles

were associated with poor metabolizer phenotype, containing four types of mutations: insertion, deletion, duplication and base substitution. Additional screening step was performed to identify CYP2D6 alleles that contain exclusively missense mutations, selecting only two alleles: CYP2D6*7 and CYP2D6*14A.

The mutations identified were validated through dbSNP and Ensembl databases using the follow criteria: multiple observations (cluster), validation in 1000 Genomes Project and presence of mutation frequency (MAF). Since CYPalleles and dbSNP use different assemblies (M333.88 and GRCh38, respectively), the Leiden Open Variation Database platform (LOVD database – [22]) was accessed to correlate data from both and to provide additional genomic information about polymorphisms position.

Functional & structural impact of missense mutations

A total of seven algorithms were used to evaluate the impact of each mutation on CYP2D6 structure and function (Table 1). To determine whether a polymorphism is deleterious or not, a threshold for each algorithm was established based on the algorithm description. Thus, a classification system was adopted as follow: 1 was attributed when the prediction indicates deleterious effect; 0 was attributed when the algorithm's prediction showed no effect of the mutation on enzyme structure or function.

The SNPeff output shows the variation of Gibbs free energy after the mutation changes in the molecule, classifying it into seven classes ranging from 'no effect in stability' to 'greatly enhanced stability' [23]. MuPro is a structural algorithm based on Support Vector Machine or Neural Network, but only Support Vector Machine results were analyzed, following the author's recommendation [24]. MutPred algorithm generates two scores: a general score (g), predicting the chance for a mutation to be deleterious or disease associated; and a property score (p) with the top five properties that may be impacted by the polymorphism [25]. SNAP2 is based on SIFT [30] and Polyphen2 [31] output data, generating a score associated with the prediction of one mutation as deleterious or not for the molecule [26]. SNPs&GO is dedicated to gene ontology analysis based on PHANTER data output [32] but also shows PhD-SNP [33] prediction as the result [27]. SNPs&GO showed to be useful for removing false positive results from data output [20]. MutationTaster2 is a Bayes classifier that provides conservation scores (Philo-P and PhastCons) and provide a general score for amino acid substitution impact. Philo-P score was used as threshold once mutations in conserved residues and regions show

Table 1. Algorithms applied to *in silico* analysis.

Algorithm	Method	Analysis	Threshold	Remarks	Ref.
SNPeffect	Naive bayes	Structural	>0.5	FoldX prediction in $\Delta\Delta G$	[23]
MuPro	SVM and neural network	Structural	<-0.5	SVM output is more accurate	[24]
MutPred	Random forest	Structural/functional	$g > 0.75$	High sensibility and specificity	[25]
SNAP2	Neural network	Functional	>50 score	Dataset construction and training included a specific data for enzymes	[26]
SNPs&GO	SVM	Functional	>0.5; RI ≥ 5	Utilizes gene ontology and is indicated for identifying false positives	[27]
Mutation Taster2	Bayes classifier	Functional	1	Provides conservational scores Phylo-P and PhastCons	[28]
SDM	Statistical	Structural	$\pm 2 \text{ kcal mol}^{-1}$	Analysis of multiple chains mutated	[29]

Δ : Variation; RI: Reliability index; SVM: Support sector machine.

a higher probability to impact the protein structure and function [28].

The Site Direct Mutator (SDM) algorithm was used to analyze the structural impact generated by accumulative mutations in the *CYP2D6* tetramer. The variants were generated by incrementing each mutation in each monomer following the chain organization. SDM uses statistical potential energy function that predicts the effect of missenses mutation in the stability of proteins, and its output includes side-chain solvent accessibility that ranges from inaccessible (<17%) and partially accessible (17–43%) to accessible (>47%). It also gives a score of pseudo $\Delta\Delta G$ that measures the variation of free energy between unfolded and folded protein, thus, giving a prediction of protein stability and disease association in protein variants [29]. SDM do not distinguish the impact of the chain affected according to molecule assembly, only the cumulative effect according to the number of chains mutated in the 3D model.

Tridimensional analysis

The *CYP2D6* crystallographic structure was used for structural prediction. 2F9Q file (Protein Database – PDB) corresponds to *CYP2D6* tetramer of four identical chains (A–D) containing 479 residues in each one, with 3.0 Å resolution. This 3D structure initiates at residue 34 and lacks the residues 42–51, but is the only one available for *CYP2D6* without substrate or inhibitor. PyMOL (The PyMOL Molecular Graphics System, Version 1.7.4, Schrödinger, LLC, NY, USA) was used to observe the 3D structure of *CYP2D6* and

also to generate the variants models to analyze the cumulative effect of each mutation through SDM.

Results

*CYP2D6*7* poor metabolizer showed only one missense mutation, described as the cause of enzyme inactivation. The polymorphism 2935A>C, present only in this allele, occurs in exon 6 resulting in H324P mutation. *CYP2D6*14A* allele shows the same phenotype caused by four missenses mutations, each one in a different exon. R296C and S486T are well distributed in all metabolizer phenotypes; while G169R is present only in *CYP2D6*14A*. P34S mutation is present in both intermediate and poor metabolizers phenotypes.

Figure 1 represents the *CYP2D6* 3D structure, showing the catalytic site and the point of each mutation for *CYP2D6*7* and *CYP2D6*14A*. Table 2 shows the algorithms prediction output for the effect of each mutation in *CYP2D6*7* and *CYP2D6*14A* over protein stability and function. Table 3 shows the SDM data output related to each mutation and the cumulative effect of the mutation in more than one chain.

Allele *CYP2D6*7*

The substitution of a polar for a nonpolar amino acid leads to an extreme change in the hydrophobicity degree, ranging from 0.165 for Histidine to 0.711 for Proline. For H324P mutation, the structural algorithms SNPeffect and MuPro predict a decrease in enzyme stability. SDM predictions show a great loss in solvent accessibility when Proline is replaced by Histidine, changing from partially accessible to a

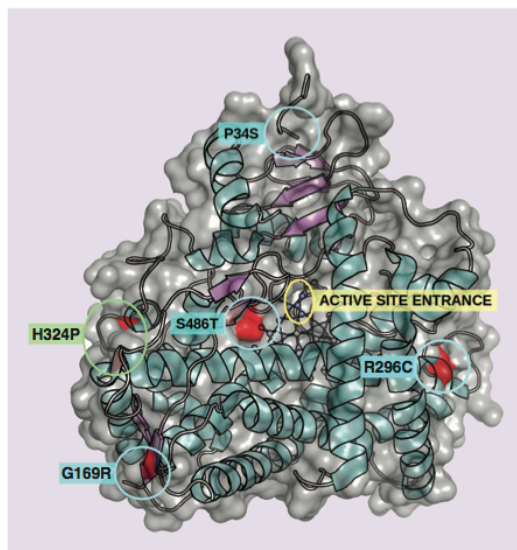


Figure 1. CYP2D6 3D structure showing the active site entrance (yellow) with the heme group; the H324P mutation carried by *CYP2D6*7* (green) and the set of mutations for the allele *CYP2D6*14A* (blue).

buried condition in the molecule, with around 22% overall reduction. The presence of H324P in only one CYP2D6 chain leads to a molecular destabilization, but its accumulation in other chains increases this impact. SDM prediction is related to the decrease of pseudo $\Delta\Delta G$, leading to an increased deleterious effect that could be associated with the enzyme dysfunction.

SNPs&GO, SNAP2 and MutPred, all algorithms for functional prediction, also showed deleterious effect of H324P. MutationTaster2 analysis showed association with disease due to the functional impact; with Phylo-P conservation full score 1.00 indicating this mutation occurs in a conserved region, leading to high functional impact.

The Proline is an imino acid that imposes severe restriction on angles accessible to the pre-proline residues in the sequence. The CYP2D6 present another

proline in position 325, therefore, this mutation occurs in a two residues junction region connecting J-helix and I-helix. The I-helix contains six residues (L302, A305, G306, T309, T310 and T313) composing the P450 active site pocket (Figure 2). So, this new restriction certainly reduces the states accessible to the backbone for 324-'PP'-325, consequently, restricting the relative position freedom of I-helix and J-helix.

Allele CYP2D6*14A

P34S mutation showed no effect on CYP2D6 stability according to SNPeffect while MuPro prediction showed an increase of CYP2D6 structural stability (Supplementary Table 1). SDM also predicted mutation effect, due to the increase of pseudo $\Delta\Delta G$, but only when P34S occurs in two or more chains. This mutation results in additional H-bond between the Ser34 in the loop and Ser70 in a β -sheet when Proline is substituted by the Serine (Figure 3). Phylo-P conservation output showed the highest score, indicating P34S functional impact prediction.

CYP2D6*14A G169R leads to a notable reduction in hydrophobicity degree once nonpolar glycine (0.501) is substituted for a polar arginine (0.000). SNPeffect and MuPro predicted a decrease in structural stability. SDM predicted an increase in solvent accessibility in 53%, with an increase in molecule destabilizing. Disease association occurs only when G169R is present in two or more chains. All functional algorithms predict damage, with a very confident hypothesis of gain of solvent accessibility ($p = 0.0037$) according to MutPred. G169R mutation shows the great importance for CYP2D6 stability, probably due to its presence in a monomer interface region (Figure 4), which could difficult the tetramer assembly.

CYP2D6*14A R296C polymorphism showed no effect on molecule structure according to SNPeffect while MuPro predicted the destabilizing effect of this mutation. SDM predicts high stabilizing effect with disease association even when the mutation is present

Table 2. Algorithms predictions for missenses mutations in *CYP2D6*7* and *CYP2D6*14A* alleles.

ID	Allele	Exon	AA change	CYPalleles (M333.88)	SNPeffect	MuPro	MutPred	SNAP2	SNPs&GO	Mutation Taster2 (Phylo-P)
rs5030867	*7	6	H324P	2935A>C	1	1	1	1	1	1
rs1065852	*14A	1	P34S	100C>T	0	1	1	1	1	1
rs5030865	*14A	3	G169R	1758G>A	1	1	1	1	1	0
rs16947	*14A	6	R296C	2850C>T	0	1	0	0	0	ND
rs1135840	*14A	9	S486T	4180G>C	0	0	0	0	0	ND

Deleterious predictions are classified as 1, and null variant are classified as 0, following the established threshold.
ND: Not determined.

Table 3. SDM outputs for amino acids solvent accessibility is showed in percentage. A is for accessible; P for partially accessible and B is for buried or inaccessible. The prediction for disease association is represented as 1 for disease-associated, and 0 for non-disease associated, following the established threshold.

Missenses mutations	Number of mutated chains	Amino acid solvent accessibility	Amino acid Δ solvent accessibility (%)	Pseudo $\Delta\Delta G$	Stability	Prediction	Disease
H324P	Wild-type	35.95% (P)					
	1	13.9% (B)	-21.9	-1.68	Destabilizing		0
	2	12.8% (B)	-23.7	-3.35	Highly destabilizing		1
	3	12.1% (B)	-24	-5.03	Highly destabilizing		1
P34S	4	13.4% (B)	-22	-6.71	Highly destabilizing		1
	Wild-type	25.17% (P)					
	1	16.6% (B)	-7.6	1.57	Stabilizing		0
	2	15.2% (B)	-6.7	2.07	Highly stabilizing		1
G169R	3	18.1% (B)	-8.6	9.96	Highly stabilizing		1
	4	17.3% (B)	-3.4	9.78	Highly stabilizing		1
	Wild-type	13.85% (B)					
	1	66.1% (A)	53	-1.82	Destabilizing		0
R296C	2	72.5% (A)	57.7	-3.65	Highly destabilizing		1
	3	66.4% (A)	51.5	-5.47	Highly destabilizing		1
	4	71.9% (A)	59.3	-7.3	Highly destabilizing		1
	Wild-type	27.55% (P)					
S486T	1	12.1% (B)	-15.6	2.47	Highly stabilizing		1
	2	11.9% (B)	-16.8	4.93	Highly stabilizing		1
	3	11.7% (B)	-14.4	7.4	Highly stabilizing		1
	4	11.5% (B)	-16.8	9.86	Highly stabilizing		1
S486T	Wild-type	52.37% (A)					
	1	47.4% (P)	-4	0.97	Slightly stabilizing		0
	2	54.3% (P)	1	1.94	Stabilizing		0
	3	47.8% (P)	-4	2.91	Highly stabilizing		1
S486T	4	55.7% (P)	2.7	3.88	Highly stabilizing		1

SDM outputs for amino acids solvent accessibility is showed in percentage. The prediction for disease association is represented as 1 for disease-associated, and 0 for non-disease associated, following the established threshold.
 Δ : Variation; A: Accessible; B: Buried or inaccessible; P: Partially accessible.

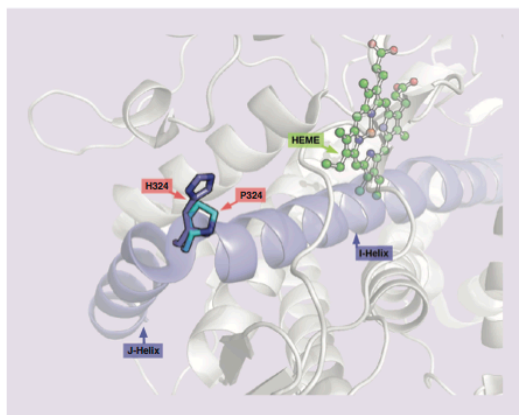


Figure 2. Polymorphism H324P from *CYP2D6*7* found between helices J and I. The wild residue histidine is presented in dark blue, the mutated amino acid proline appears in light blue. Heme group is represented in green near to I-helix extremity.

in only one chain. It may appear counterintuitive that increased protein stability can lead to protein malfunction. However, according to SDM authors, protein flexibility is essential for enzyme catalysis. MutPred also predicted R296C as low probability of causing loss of enzyme function but pointed an actionable hypothesis to cause loss of disorder ($p = 0.0411$). SNPs&GO and SNAP2 showed no effect of R296C over protein function.

*CYP2D6*14A* S486T shows no structural damage according to SNPeff. Despite MuPro predicted the increase of structural stability; the result has low confidence score (Supplementary Table 1). SDM output shows disease association, but only if the mutation occurs in at least three chains, increasing structure stabilization due to S486T accumulation. No functional effect was found for this mutation.

Discussion

CYP2D6 is a clinically important enzyme that metabolizes numerous therapeutic drugs. The enzyme is responsible for the oxidative metabolism of up to 25% of commonly prescribed drugs such as antidepressants, antipsychotics, opioids, antiarrhythmics and tamoxifen. The ultra-rapid metabolizer phenotype is recognized as a cause of therapeutic inefficacy of antidepressant, whereas an increased risk of toxicity has been reported in poor metabolizers with several psychotropics. However, conflicting results have been reported regarding the association between *CYP2D6* genotype and tamoxifen effects. In this way, the *CYP2D6* genotyping may be useful in selecting an adjuvant hormonal therapy in postmenopausal women [34]. The present study provides *in silico* analysis regarding functional alterations of *CYP2D6*7* and *CYP2D6*14A* vari-

ant proteins. To assess the effects of *CYP2D6* variant allele's, algorithms were applied to the wild and mutant sequence/structure.

*CYP2D6*7* was first described by Evert *et al.* (1994), so-called mutation 'E.' The author suggested that H324P occurs near to the active site and also associates this mutation to a complete loss of enzyme function [35]. Later, the same author performed tests with recombinant DNA proving the null activity of H324P variant once it prevents the correct folding and the incorporation of heme group to CYP2D6 [36]. The expression of H324P mutation by recombinant baculovirus results in a protein that was practically insoluble, thus demonstrating that it also strongly influences subcellular localization of recombinant synthesized *CYP2D6*7* [36]. In our analysis, this structural/functional effect of H324P showed to be due to mutation occurrence in a high-conserved amino acid. This fact leads to the reduction of flexibility in the α -helix involved in the stabilization of heme group, extending its impact to the active site dynamics.

The effect of H324P over the molecule seems to be so relevant that it is the only mutation described for *CYP2D6*7*. By another side, *CYP2D6*7* seems to be either absent or rare allele, according to the population studied. Among 112 Hungarian individuals the frequency for *CYP2D6*7* was 0.4% [37], while among 165 Caucasian women of Geneva the frequency was

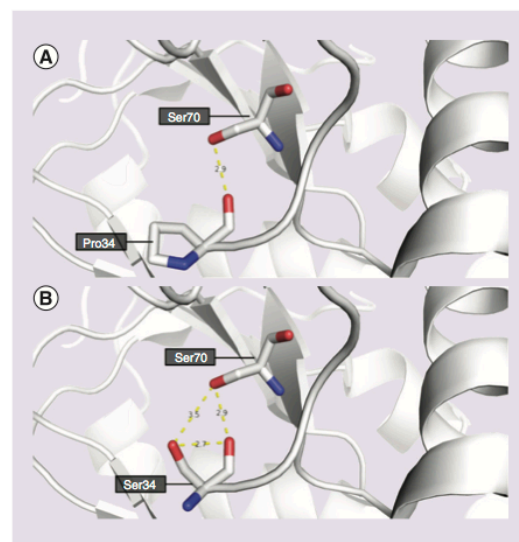


Figure 3. 3D representation of mutated CYP2D6 protein with P34S polymorphism from *CYP2D6*14A* allele. In Figure 3A is presented the wild Proline-34 residue making an H-bond (yellow) with Serine-70. Figure 3B shows the mutated residue 34 with Serine substitution and the additional H-bond added upon polymorphism. Supplementary Table 1: Algorithms original outputs without any threshold applied.

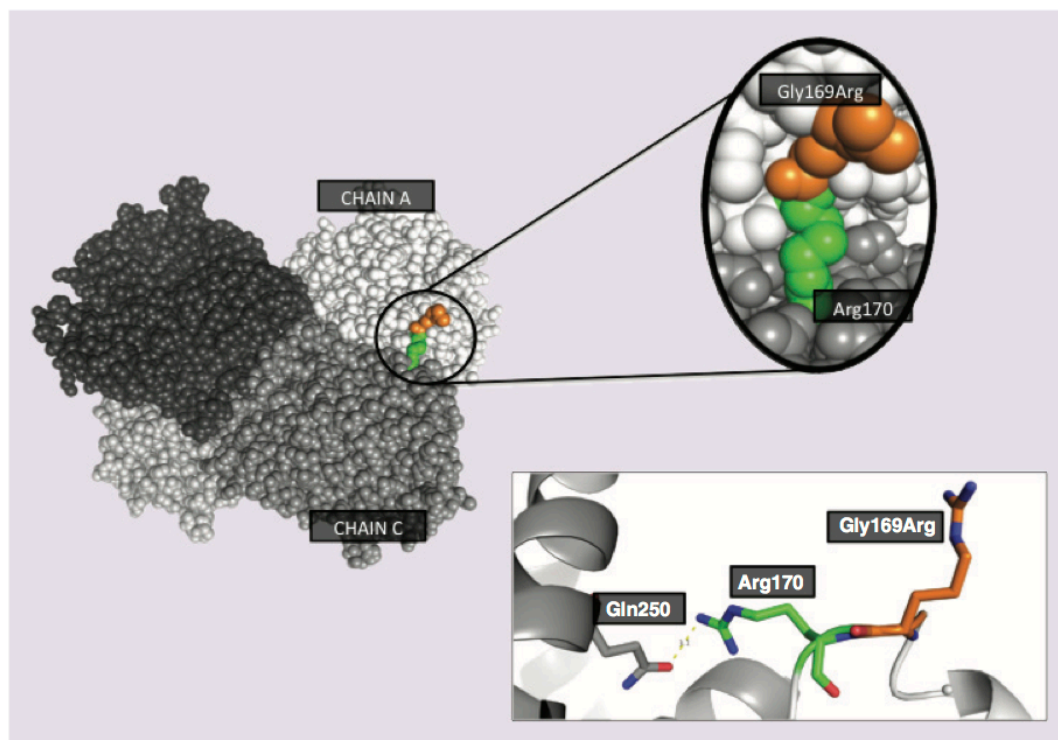


Figure 4. 3D representation of mutated *CYP2D6* protein with G169R polymorphism from *CYP2D6*14A* allele. The replacement of a glycine for an arginine (orange) in position 169 occurs in the interface region between chains A and C. The Arginine-170 (green) in chain A, contributes to tetramer stabilization through an H-bond formation with Glutamine-250 from chain C.

0.3% [38]. No frequency was observed in 264 Mexican American [39], 105 Spanish [40] and 323 German Caucasian [41] individuals. Since there are few reports about *CYP2D6*7*, the frequency in the multiracial population is not well reported neither its influence in breast cancer treatment.

The *CYP2D6*14* allele has two subtypes. *CYP2D6*14A* carries the P34S, G169R, R296C and S486T substitutions and reports only detectable activity while *CYP2D6*14B* carries G169R, R296C and S486T missenses substitutions, which results in high enzyme activity. These data suggest that the relatively high activity of *CYP2D6*14B* may be attributed mainly to the absence of the P34S substitution [42]. Regarding this mutation, the presence of Serine tends to reverse the enhanced flexibility, given that it removes the typical proline restrictions on backbone angles.

According to SDM, P34S increases the stability of the enzyme in such a way that causes loss of function. Experimental analysis showed that the decrement in metabolic activity by G169R mutation is not dramatic as P34S mutation [43]. Information theory-based analysis also predicted G169R to be leaky and to be consistent with decreased (but not totally abolished)

mRNA production, once its presence at the end of exon 3 reduces splice site use at least eightfold [44]. R296C mutation generates a loss disorder, despite no great deleterious effect can be observed.

This result was also observed in cloning experiments in *Escherichia coli* resulting in <10% of enzyme activity [45]. However, experimental tests with eukaryotic cells, which have a better protein folding, could be performed to understand its influence in the enzyme active site. By another side, cloning experiments for S486T in eukaryotic cells showed no consequence for *CYP2D6* function [46], as we observed in our analysis.

*CYP2D6*14* allele showed to be more frequent in the Asian population, with 1.4% in 295 Chinese individuals [47] and 2.2% in 162 Japanese [48]. However, among 90 individual from Sri Lanka (Sinhalese, Tamils and Moors), no *CYP2D6*14* was found [49]. No frequency was observed in Mexican–American [39], Spanish [40] and Saudi Arabian populations [50]. The combination of G169R with P34S, R296C and S486T mutations may have further diminished the metabolic activity of *CYP2D6* [43], and would be very interesting to perform *in silico* tests for the possible additive effect of these four mutation sites in the same molecule.

Conclusion

In silico analysis have shown evidence that missense mutations in *CYP2D6**7 and *CYP2D6**14A cause *CYP2D6* dysfunction. We demonstrated that conformational 3D alterations were responsible for the *CYP2D6* malfunction that leads to tamoxifen under-performance. Further analysis should be performed to all *CYP2D6* alleles associated with poor metabolizer phenotype to improve the pharmacogenomics application in breast cancer therapy.

Future perspective

CYP2D6 is an emerging biomarker that has been extensively investigated, but current guidelines do not recommend routine genetic testing of it [51]. Even so, determining whether a patient will or not benefit from tamoxifen therapy is of paramount importance in the clinical course of cancer treatment. The ROCHE AmpliChip is the closest technology available to *CYP2D6* genotyping [38], but the cost is still too high that its use is limited. To allow the development of new devices for fast, precise and low-cost detection of *CYP2D6* polymorphisms is necessary to recognize the most important ones regarding poor metabolization phenotype. Toward to the near future, *CYP2D6*

genotyping will be part of breast cancer routine treatment, and then new point-of-care devices will be used in hospitals and outpatient to guide and monitor the treatment.

Supplementary data

To view the supplementary data that accompany this paper, please visit the journal website at: www.futuremedicine.com/doi/full/10.2217/psg-2015-0003

Acknowledgements

The authors thank AM da Silva Neto for the contributions in the discussion and figures design. They also thank P Radivojac and coworkers for their support and availability in elucidating MutPred results.

Financial & competing interests disclosure

This work was supported by Pernambuco State Science and Technology Support Foundation (FACEPE). The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

No writing assistance was utilized in the production of this manuscript.

Executive summary

Background

- *CYP2D6* polymorphisms may be the cause of tamoxifen failure in breast cancer treatment.
- *CYP2D6* is a high polymorphic enzyme from P450 cytochrome, responsible for metabolizing almost 25% of drugs.
- Tamoxifen is an antineoplastic prodrug for estrogen receptor positive breast cancer that depends on *CYP2D6* to be metabolized and generate the active metabolites.
- Almost 50% of breast cancer patients do not benefit from tamoxifen treatment.

Methods

- *In silico* analysis were performed over *CYP2D6**7 and *CYP2D6**14A poor metabolizer alleles mutations to access the impact of missense mutation on protein stability and function.

Results

- The allele *CYP2D6**7 have only one mutation described; H324P predictions showed high deleterious impact on *CYP2D6* structure and function.
- The allele *CYP2D6**14A has four missense mutations, and three of them are associated with loss of enzyme stability and function.

Conclusion

- This study was the first one to predict the structural impact of the missense mutations on *CYP2D6* catalysis. This information supports the hypothesis that *CYP2D6* can be a good marker for tamoxifen pharmacogenomics.

References

- 1 Ferlay J, Soerjomataram I, Ervik M *et al*. Breast cancer estimated incidence, mortality and prevalence worldwide in 2012. <http://globocan.iarc.fr>
- 2 Lumachi F, Santeufemia DA, Basso SM. Current medical treatment of estrogen receptor-positive breast cancer. *World J. Biol. Chem.* 6(3), 231–239 (2015).
- 3 Groenendijk FH, Bernards R. Drug resistance to targeted therapies: déjà vu all over again. *Mol. Oncol.* 8(6), 1067–1083 (2014).
- 4 Binkhorst L, Mathijssen RHJ, Jager A, van Gelder T. Individualization of tamoxifen therapy: much more than just *CYP2D6* genotyping. *Cancer Treat. Rev.* 41(3), 289–299 (2015).
- 5 Klein DJ, Thorn CF, Desta Z, Flockhart DA, Altman

Evaluating the impact of missenses mutations in CYP2D6*7 & CYP2D6*14A Short communication

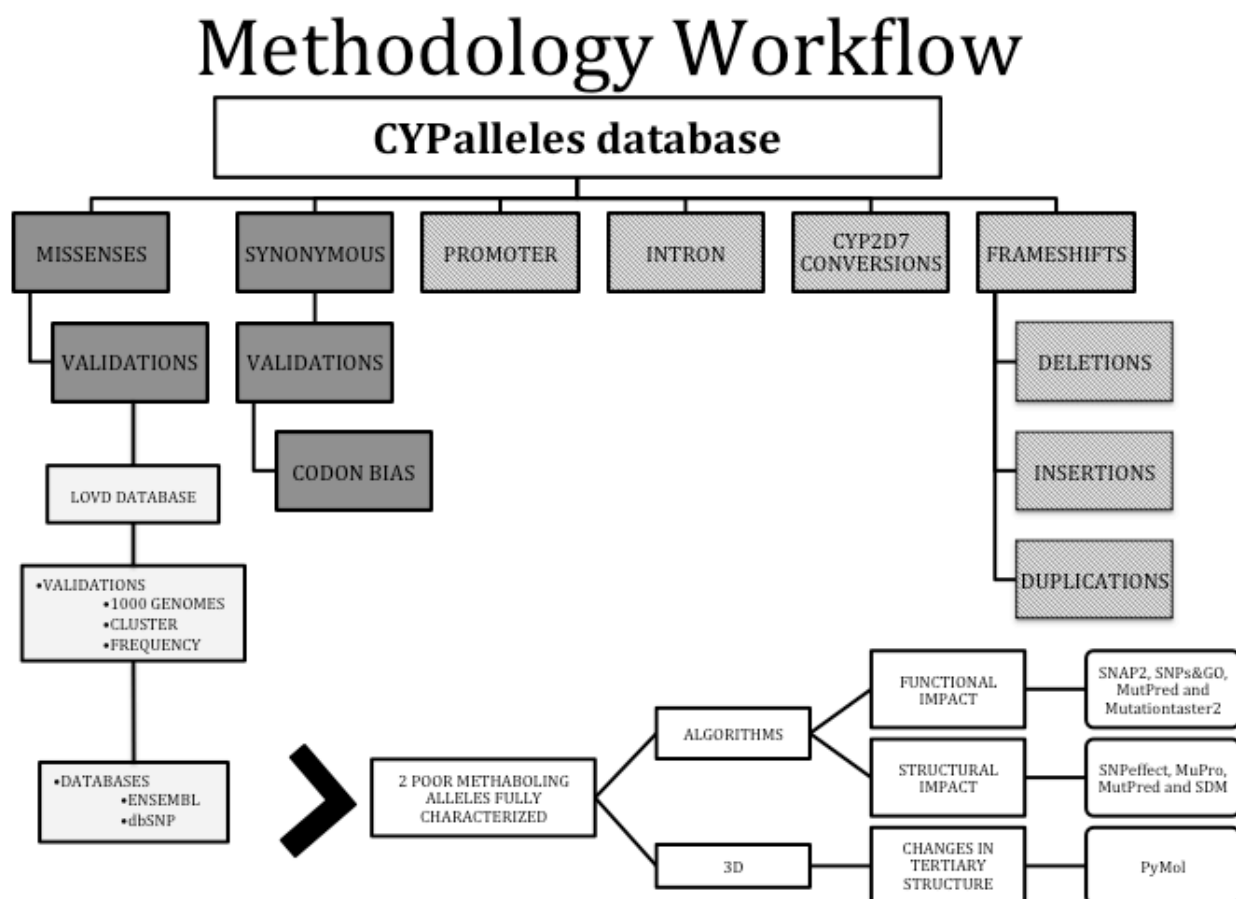
- RB, Klein TE. PharmGKB summary: tamoxifen pathway, pharmacokinetics. *Pharmacogenet. Genomics* 23(11), 1199–1216 (2013).
- 6 Zanger UM, Schwab M. Cytochrome P450 enzymes in drug metabolism: regulation of gene expression, enzyme activities, and impact of genetic variation. *Pharmacol. Ther.* 138(1), 103–141 (2013).
 - 7 Rowland P, Blaney F, Smyth M *et al.* Crystal structure of human cytochrome P450 2D6. *J. Biol. Chem.* 281(11), 7614–7622 (2006).
 - 8 Hertz DL, Snavelly AC, McLeod HL *et al.* *In vivo* assessment of the metabolic activity of CYP2D6 diplotypes and alleles. *Br. J. Clin. Pharmacol.* 80(5), 1122–1130 (2015).
 - 9 The Human Cytochrome P450 (CYP) Allele Nomenclature Database. www.cypalleles.ki.se/cyp2d6.htm
 - 10 Chin FW, Chan SC, Abdul Rahman S, Noor Akmal S, Rosli R. CYP2D6 genetic polymorphisms and phenotypes in different ethnicities of Malaysian breast cancer patients. *Breast J.* 22(1), 54–62 (2015).
 - 11 Province M, Goetz MP, Brauch H *et al.* CYP2D6 genotype and adjuvant tamoxifen: meta-analysis of heterogeneous study populations. *Clin. Pharmacol. Ther.* 95(2), 216–227 (2014).
 - 12 Goetz MP, Suman VJ, Hoskin TL *et al.* CYP2D6 metabolism and patient outcome in the Austrian Breast and Colorectal Cancer Study Group trial (ABCSG) 8. *Clin. Cancer Res.* 19(2), 500–507 (2013).
 - 13 Schroth W, Goetz MP, Hamann U *et al.* Association between CYP2D6 polymorphisms and outcomes among women with early stage breast cancer treated with tamoxifen. *JAMA* 302(13), 1429–1436 (2014).
 - 14 Regan MM, Leyland-Jones B, Bouzyk M *et al.* CYP2D6 genotype and tamoxifen response in postmenopausal women with endocrine-responsive breast cancer: The Breast International Group 1–98 trial. *J. Natl. Cancer Inst.* 104(6), 441–451 (2012).
 - 15 Rae JM, Drury S, Hayes DF *et al.* CYP2D6 and UGT2B7 genotype and risk of recurrence in tamoxifen-treated breast cancer patients. *J. Natl. Cancer Inst.* 104(6), 452–60 (2012).
 - 16 Ruddy KJ, Desantis SD, Gelman RS *et al.* Personalized medicine in breast cancer: tamoxifen, endoxifen, and CYP2D6 in clinical practice. *Breast Cancer Res. Treat.* 141, 421–427 (2013).
 - 17 Cannataro M, Santos RW Dos, Sundnes J. Biomedical and bioinformatics challenges to computer science: bioinformatics, modeling of biomedical systems and clinical applications. *Procedia Comput. Sci.* 4, 1058–1061 (2011).
 - 18 Bellazzi R, Masseroli M, Murphy S, Shabo A, Romano P. Clinical bioinformatics: challenges and opportunities. *BMC Bioinformatics* 13(S14), 1471–2105 (2012).
 - 19 Dong C, Wei P, Jian X *et al.* Comparison and integration of deleteriousness prediction methods for nonsynonymous SNVs in whole exome sequencing studies. *Hum. Mol. Genet.* 24(8), 2125–2137 (2015).
 - 20 Frousios K, Iliopoulos CS, Schlitt T, Simpson M. Predicting the functional consequences of non-synonymous DNA sequence variants – evaluation of bioinformatics tools and development of a consensus strategy. *Genomics* 102(4), 223–228 (2013).
 - 21 Fiser A. Template-based protein structure modeling. *Methods Mol. Biol.* 673, 73–94 (2010).
 - 22 Leiden Open Variation Database. www.lovd.nl/3.0/home
 - 23 De Baets G, Van Durme J, Reumers J *et al.* SNPeffect 4.0: on-line prediction of molecular and structural effects of protein-coding variants. *Nucleic Acids Res.* 40(D1), D935–D939 (2012).
 - 24 Cheng J, Randall A, Baldi P. Prediction of protein stability changes for single-site mutations using support vector machines. *Proteins* 62(4), 1125–1132 (2006).
 - 25 Li B, Krishnan VG, Mort ME *et al.* Automated inference of molecular mechanisms of disease from amino acid substitutions. *Bioinformatics* 25(21), 2744–2750 (2009).
 - 26 Hecht M, Bromberg Y, Rost B. Better prediction of functional effects for sequence variants. *BMC Genomics*. 16(Suppl 8), S1 (2015).
 - 27 Calabrese R, Capriotti E, Fariselli P, Martelli PL, Casadio R. Functional annotations improve the predictive score of human disease-related mutations in proteins. *Hum. Mutat.* 30(8), 1237–1244 (2009).
 - 28 Schwarz JM, Cooper DN, Schuelke M, Seelow D. MutationTaster2: mutation prediction for the deep-sequencing age. *Nat. Methods*. 11(4), 361–362 (2014).
 - 29 Worth CL, Preissner R, Blundell TL. SDM – a server for predicting effects of mutations on protein stability and malfunction. *Nucleic Acids Res.* 39(Web Server issue), 215–222 (2011).
 - 30 Ng PC, Henikoff S. Predicting deleterious amino acid substitutions. *Genome Res.* 11, 863–874 (2001).
 - 31 Adzhubei Ia, Schmidt S, Peshkin L *et al.* A method and server for predicting damaging missense mutations. *Nat. Methods* 7(4), 248–249 (2010).
 - 32 Thomas PD, Kejariwal A, Guo N *et al.* Applications for protein sequence-function evolution data: mRNA/protein expression analysis and coding SNP scoring tools. *Nucleic Acids Res.* 34(Web Server issue), 645–650 (2006).
 - 33 Capriotti E, Calabrese R, Casadio R. Predicting the insurgence of human genetic diseases associated to single point protein mutations with support vector machines and evolutionary information. *Bioinformatics* 22(22), 2729–2734 (2006).
 - 34 Samer CF, Lorenzini KI, Rollason V, Daali Y, Desmeules JA. Applications of CYP450 testing in the clinical setting. *Mol. Diagn. Ther.* 17(3), 165–184 (2013).
 - 35 Evert B, Gries E, Eichebaum M. A missense mutation in exon 6 of the CYP2D6 gene leading to a histidine 324 to proline exchange is associated with the poor metabolizer phenotype of sparteine. *Naunyn Schmiedebergs Arch. Pharmacol.* 350, 434–439 (1994).
 - 36 Evert B, Eichelbaum M, Haubruck H, Zanger UM. Functional properties of CYP2D6 1 (wild-type) and CYP2D6 7 (His324Pro) expressed by recombinant baculovirus in insect cells. *Naunyn Schmiedebergs Arch. Pharmacol.* 355(3), 309–318 (1997).

Short communication Borba, Melo-Neto, Leitão, Castelletti, Lima-Filho & Martins

- 37 Rideg O, Háber A, Botz L *et al.* Pilot study for the characterization of pharmacogenetically relevant *CYP2D6*, *CYP2C19* and *ABCB1* gene polymorphisms in the Hungarian population. *Cell Biochem. Funct.* 29(7), 562–568 (2011).
- 38 Rebsamen MC, Desmeules J, Daali Y *et al.* The AmpliChip CYP450 test: cytochrome P450 2D6 genotype assessment and phenotype prediction. *Pharmacogenomics J.* 9(1), 34–41 (2009).
- 39 Luo H-R, Gaedigk A, Aloumanis V, Wan Y-JY. Identification of CYP2D6 impaired functional alleles in Mexican Americans. *Eur. J. Clin. Pharmacol.* 61(11), 797–802 (2005).
- 40 Menoyo A, del Rio E, Baiget M. Characterization of variant alleles of cytochrome *CYP2D6* in a Spanish population. *Cell Biochem. Funct.* 24(5), 381–385 (2006).
- 41 Stamer UM, Bayerer B, Wolf S, Hoeft A, Stüber F. Rapid and reliable method for cytochrome P450 2D6 genotyping. *Clin. Chem.* 48, 1412–1417 (2002).
- 42 Sakuyama K, Sasaki T, Ujiie S *et al.* Functional characterization of 17 *CYP2D6* allelic variants (CYP2D6.2, 10, 14A-B, 18, 27, 36, 39, 47–51, 53–55, and 57). *Pharmacology.* 36(12), 2460–2467 (2008).
- 43 Wang SL, Lai MD, Huang JD. G169R mutation diminishes the metabolic activity of CYP2D6 in Chinese. *Drug Metab. Dispos.* 27(3), 385–388 (1999).
- 44 Rogan PK, Svojanovsky S, Leeder JS. Information theory-based analysis of CYP2C19, CYP2D6 and CYP3A5 splicing mutations. *Pharmacogenetics* 13(4), 207–218 (2003).
- 45 Kim J, Lim Y-R, Han S *et al.* Functional influence of human *CYP2D6* allelic variations: P34S, E418K, S486T, and R296C. *Arch. Pharm. Res.* 36(12), 1500–1506 (2013).
- 46 Kagimoto M, Heim M, Kagimoto K, Zeugin T, Meyer U. Multiple mutations of the human cytochrome P450IID6 (*CYP2D6*) in poor metabolizers of debrisoquine. *J. Biol. Chem.* 265(26), 17209–17214 (1990).
- 47 Cai WM, Chen B, Zhang WX. Frequency of *CYP2D6**10 and *14 alleles and their influence on the metabolic activity of CYP2D6 in a healthy Chinese population. *Clin. Pharmacol. Ther.* 81(1), 95–98 (2007).
- 48 Nishida Y, Fukuda T, Yamamoto I, Azuma J. CYP2D6 genotypes in a Japanese population: low frequencies of *CYP2D6* gene duplication but high frequency of *CYP2D6**10. *Pharmacogenetics* 10(6), 567–570 (2000).
- 49 Tharanga TD, Jinadasa CM, Risama MF, Galappaththy P, Jayakody RL, Dissanayake VH. Genetic variants in the cytochrome P450 2D6 gene in the Sri Lankan population. *Indian J. Hum. Genet.* 19(4), 392–396 (2013).
- 50 Al-Dosari MS, Al-Jenoobi FI, Alkharfy KM *et al.* High prevalence of *CYP2D6**41 (G2988A) allele in Saudi Arabians. *Environ. Toxicol. Pharmacol.* 36(3), 1063–1067 (2013).
- 51 Kalia M. Biomarkers for personalized oncology: recent advances and future challenges. *Metabolism* 64(3), S16–S21 (2015).

4.1 Material Suplementar

Supplementary Figure 1



Supplementary Table 1: Algorithms original outputs without any threshold applied.

SNPEffect			MuPro		MutPred	SNAP2			SNP&GO output			MUTATION TASTER2		
$\Delta\Delta G$ (kcal/mol)	TotX prediction	Effect	Confidence score	Score	Hypothesis	Prediction	Score	Accuracy	PREDICTION	RI	PROBABILITY Y	Score	PhastCons	PhyloP
H324P	5.75	Decrease	-0.67388092	0.781	Confident: Gain of catalytic residue at H324 (P = 0.0186)	Effect	79	85%	Disease	5	0.772	77	2.854	1
P34S	0.38	Null	-0.98278866	0.752	Actionable: Loss of catalytic residue at P34 (P=0.0091), Loss of glycosylation at P34 (P = 0.0373), Gain of loop (P = 0.0435)	Effect	66	80%	Disease	0	0.515	74	4.426	1
G169R	1.84	Reduction	-0.11462576	0.803	Very Confident: Gain of solvent accessibility (P = 0.0037)	Effect	27	63%	Disease	5	0.766	125	0.797	0.419
R296C	-0.31	Null	-0.95625891	0.595	Actionable: Loss of disorder (P = 0.0411)	Neutral	-69	82%	Neutral	7	0.157	ND	ND	ND
S486T	0.42	Null	0.17381188	0.291	-	Neutral	-98	97%	Neutral	9	0.033	ND	ND	ND

5 CONCLUSÕES E PERSPECTIVAS

Foram identificados dois alelos pobre metabolizadores da CYP2D6 contendo apenas mutações missenses. Os alelos CYP2D6*7 e CYP2D6*14A contêm juntos cinco polimorfismos que foram submetidos a análise através de sete algoritmos de predição de impacto estrutural e funcional. Posteriormente a avaliação 3D foram demonstradas as consequências dos polimorfismos a nível de estrutura terciária. Os resultados apresentados neste trabalho forneceram evidências importantes para a elucidação do papel dos polimorfismos da CYP2D6 na farmacogenética do câncer de mama, gerando publicação em revista internacional de alto impacto.

As principais perspectivas deste estudo incluem realizar análise do impacto de múltiplas mutações em uma mesma proteína, a fim de mimetizar o fenótipo real de um paciente pobre metabolizador; e realizar análises de interação entre as proteínas mutadas e o tamoxifeno, para observar se os polimorfismos reduzem a afinidade da CYP2D6 pelo seu substrato.

REFERÊNCIAS

- ALENCAR, S. A.; LOPES, J. C. D. A comprehensive in silico analysis of the functional and structural impact of SNPs in the IGF1R gene. **Journal of biomedicine & biotechnology**, 2010.
- ANTUNES, M. V et al. Endoxifen levels and its association with CYP2D6 genotype and phenotype: Evaluation of a Southern Brazilian population under tamoxifen pharmacotherapy. **Therapeutic drug monitoring**, v. 34, n. 4, p. 422–431, 2012.
- BADVE, S. et al. Basal-like and triple-negative breast cancers: a critical review with an emphasis on the implications for pathologists and oncologists. **Modern Pathology**, v. 24, n. 2, p. 157–167, 2011.
- BAUMANN, P. **CYP2D6: Genetics, Pharmacology and Clinical Relevance**. London: Future Medicine, 2015. v. 1
- BERSE, B.; LYNCH, J. A. Molecular diagnostic testing in breast cancer. **Seminars in oncology nursing**, v. 31, n. 2, p. 108–21, 2015.
- BINKHORST, L. et al. Individualization of tamoxifen therapy: much more than just CYP2D6 genotyping. **Cancer treatment reviews**, v. 41, n. 3, p. 289–99, 2015.
- CAO, J.; XIONG, L. Protein Sequence Classification with Improved Extreme Learning Machine Algorithms. v. 2014, 2014.
- CAPRIOTTI, E. et al. Bioinformatics for personal genome interpretation. **Briefings in Bioinformatics**, v. 13, n. 4, p. 495–512, 2012.
- CHANG, J. Reward bioinformaticians. **Nature**, v. 520, n. 7546, p. 151–152, 2015.
- CHOONG, Y. S.; TYE, G. J.; LIM, T. S. Minireview: Applied structural bioinformatics in proteomics. **Protein Journal**, v. 32, n. 7, p. 505–511, 2013.
- DALY, A. K. et al. Nomenclature for human CYP2D6 alleles. **Pharmacogenetics**, v. 6, p. 193–201, 1996.
- EROLE, P. et al. Molecular biology in breast cancer: intrinsic subtypes and signaling pathways. **Cancer Treatment Reviews**, v. 38, n. 6, p. 698–707, 2012.

FERLAY, J. et al. **GLOBOCAN 2012 v1.0, Cancer Incidence and Mortality Worldwide: IARC International Agency for Research on Cancer**. Lyon, France: [s.n.].

FISER, A. Template-Based Protein Structure Modeling. **Methods in Molecular Biology**, v. 673, p. 73–94, 2010.

FRANK, M. S. The essence of SNPs. **Gene**, p. 177–186, 1999.

FROUSIOS, K. et al. Predicting the functional consequences of non-synonymous DNA sequence variants - evaluation of bioinformatics tools and development of a consensus strategy. **Genomics**, v. 102, n. 4, p. 223–228, 2013.

GEORGE PRIYA DOSS, C. et al. Extrapolating the effect of deleterious nsSNPs in the binding adaptability of flavopiridol with CDK7 protein: a molecular dynamics approach. **Human Genomics**, v. 7, n. 1, p. 10, 2013.

GOETZ, M. P. et al. CYP2D6 metabolism and patient outcome in the Austrian Breast and Colorectal Cancer Study Group trial (ABCSG) 8. **Clinical Cancer Research**, v. 19, n. 2, p. 500–7, 2013.

GROENENDIJK, F. H.; BERNARDS, R. Drug resistance to targeted therapies: déjà vu all over again. **Molecular oncology**, v. 8, n. 6, p. 1067–83, 2014.

GUNN, S. R. Support Vector Machines for Classification and Regression. **University of Southampton**, 1998.

HERTZ, D. L. et al. In Vivo Assessment of the Metabolic Activity of CYP2D6 Diplotypes and Alleles. **British Journal of Clinical Pharmacology**, v. 80, n. 5, p. 1122–1130, 2015.

INGELMAN-SUNDBERG, M. et al. Human cytochrome P450 (CYP) genes: recommendations for the nomenclature of alleles. **Pharmacogenetics**, v. 10, n. 1, p. 91–93, 2000.

INGELMAN-SUNDBERG, M.; SIM, S. C. Intronic polymorphisms of cytochromes P450. **Human genomics**, v. 4, n. 6, p. 402–405, 2010.

INSUA-RODRÍGUEZ, J.; OSKARSSON, T. The Extracellular Matrix in Breast Cancer. **Advanced drug delivery reviews**, v. 97, p. 41–55, 2015.

KALIA, M. Biomarkers for personalized oncology: recent advances and future challenges.

Metabolism, v. 64, n. 3, p. S16–S21, 2015.

KANDOI, G.; ACENCIO, M. L.; LEMKE, N. Prediction of Druggable Proteins Using Machine Learning and Systems Biology: A Mini-Review. **Frontiers in Physiology**, v. 6, n. December, 2015.

KIM, J. et al. Functional influence of human CYP2D6 allelic variations: P34S, E418K, S486T, and R296C. **Archives of pharmacal research**, v. 36, n. 12, p. 1500–6, 2013.

KUMAR, A. et al. Computational SNP analysis: current approaches and future prospects. **Cell biochemistry and biophysics**, v. 68, n. 2, p. 233–9, 2014.

KUMAR, V.; ABBAS, A. K.; FAUSTO, N. Robbins and Cotran's Pathologic Basis of Disease. **The American Journal of Surgical Pathology**, v. 34, n. 1, p. 132, 2010.

LEE, I. H. Molecular Prognostic and Predictive Assays in Breast Cancer. **American Journal of Surgical Pathology: Review and Reports**, v. 21, n. 1, p. 4–10, 2016.

LEON, J. DE; SUSCE, M. T.; MURRAY-CARMICHAEL, E. The AmpliChip™ CYP450 Genotyping Test Integrating a New Clinical Tool. **Pharmacogenetics**, v. 10, n. 3, p. 135–151, 2006.

LI, B. et al. Automated inference of molecular mechanisms of disease from amino acid substitutions. **Bioinformatics**, v. 25, n. 21, p. 2744–2750, 2009.

LI, L.; WEI, D. Bioinformatics Tools for Discovery and Functional Analysis of Single Nucleotide Polymorphisms. **Springer and Shanghai Jiao Tong University Press**, v. 827, p. 287–310, 2015.

LISBETH, P. et al. Genotype and co-medication dependent CYP2D6 metabolic activity: effects on serum concentrations of aripiprazole, haloperidol, risperidone, paliperi... - PubMed - NCBI. **European journal of clinical pharmacology**, v. 72, p. 175–184, 2016.

LIU, M. et al. Comparison of random forest, support vector machine and back propagation neural network for electronic tongue data classification: Application to the recognition of orange beverage and Chinese vinegar. **Sensors and Actuators, B: Chemical**, v. 177, p. 970–980, 2013.

LLERENA, A. et al. Interethnic variability of CYP2D6 alleles and of predicted and measured

metabolic phenotypes across world populations. **Expert opinion on drug metabolism & toxicology**, v. 10, n. 11, p. 1569–83, 2014.

LUMACHI, F.; SANTEUFEMIA, D. A.; BASSO, S. M. Current medical treatment of estrogen receptor-positive breast cancer. **World Journal of Biological Chemistry**, v. 6, n. 3, p. 231–239, 2015.

MAGEE, J. A.; PISKOUNOVA, E.; MORRISON, S. J. Cancer Stem Cells: Impact, Heterogeneity, and Uncertainty. **Cancer Cell**, v. 21, n. 3, p. 283–296, 2012.

MAKKI, J. Diversity of Breast Carcinoma : Histological Subtypes and Clinical Relevance. **Clinical Medicine Insights: Pathology**, v. 8, p. 23–31, 2015.

MARTINS, D. M. F. et al. Determination of CYP2D6 *3, *4 and *10 frequency in women with breast cancer in São Luís, Brazil , and its association with prognostic factors and disease-free survival. **Brazilian Journal of Medical and Biological Research**, v. 47, n. 11, p. 1008–1015, 2014.

MAUGERI-SACCÀ, M.; DE MARIA, R. Hippo pathway and breast cancer stem cells. **Critical Reviews in Oncology/Hematology**, v. 99, p. 115–122, 2015.

MINISTÉRIO DA SAÚDE - BRASIL. **Estimativa 2016: incidência de câncer no Brasil/Instituto Nacional de Câncer José Alencar Gomes da Silva**. Rio de Janeiro, Brasil: [s.n.].

MURAWA, P. et al. Breast cancer: Actual methods of treatment and future trends. **Reports of practical oncology and radiotherapy**, v. 19, n. 3, p. 165–72, 2014.

MÜRDTER, T. E. et al. Activity levels of tamoxifen metabolites at the estrogen receptor and the impact of genetic polymorphisms of phase I and II enzymes on their concentration levels in plasma. **Clinical pharmacology and therapeutics**, v. 89, n. 5, p. 708–17, maio 2011.

O'REILLY, E. A. et al. The fate of chemoresistance in triple negative breast cancer (TNBC). **BBA Clinical**, v. 3, p. 257–275, 2015.

OSCARSON, M.; INGELMAN-SUNDBERG, M. CYPalleles: a web page for nomenclature of human cytochrome P450 alleles. **Drug metabolism and pharmacokinetics**, v. 17, n. 6, p. 491–495, 2002.

PEROU, C. M. et al. Molecular portraits of human breast tumours. **Nature**, v. 406, n. 6797, p. 747–752, 2000.

RAE, J. M. et al. CYP2D6 and UGT2B7 genotype and risk of recurrence in tamoxifen-treated breast cancer patients. **Journal of the National Cancer Institute**, v. 104, n. 6, p. 452–60, 2012.

REGAN, M. M. et al. CYP2D6 genotype and tamoxifen response in postmenopausal women with endocrine-responsive breast cancer: the breast international group 1-98 trial. **Journal of the National Cancer Institute**, v. 104, n. 6, p. 441–51, 2012.

RELLING, M. V; EVANS, W. E. Pharmacogenomics in the clinic. **Nature**, v. 526, n. 7573, p. 343–350, 2015.

SALADORES, P. et al. Tamoxifen metabolism predicts drug concentrations and outcome in premenopausal patients with early breast cancer. **The pharmacogenomics journal**, n. April, p. 1–11, 5 ago. 2014.

SCHROTH, W. et al. Association Between CYP2D6 Polymorphisms and Outcomes Among Women With Early Stage Breast Cancer Treated With Tamoxifen. **The Journal of the American Medical Association**, v. 302, n. 13, p. 1429–1436, 2014.

SEAL, A. et al. Tools, resources and databases for SNPs and indels in sequences: a review. **International Journal of Bioinformatics Research and Applications**, v. 10, n. 3, p. 264, 2014.

SHANTHI, V.; RAJASEKARAN, R.; RAMANATHAN, K. Computational Identification of Significant Missense Mutations in AKT1 Gene. **Cell Biochemistry and Biophysics**, p. 957–965, 2014.

STEIN, R. C. et al. OPTIMA prelim: a randomised feasibility study of personalised care in the treatment of women with early breast cancer. **Health Technol Assess**, v. 20, n. 10, p. 1–202, 2016.

VAN NESTE, L.; VAN CRIEKINGE, W. We are all individuals... bioinformatics in the personalized medicine era. **Cellular Oncology**, v. 38, n. 1, p. 29–37, 2015.

VARGO-GOGOLA, T.; ROSEN, J. M. Modelling breast cancer: one size does not fit all. **Nature Reviews Cancer**, v. 7, n. 9, p. 659–672, 2007.

VERLI, H. **Bioinformática da Biologia à Flexibilidade Molecular**. 1. ed. Porto Alegre: [s.n.].

VIDEIRA, M.; REIS, R. L.; BRITO, M. A. Deconstructing breast cancer cell biology and the mechanisms of multidrug resistance. **Biochimica et Biophysica Acta (BBA) - Reviews on Cancer**, v. 1846, n. 2, p. 312–325, 2014.

WANG, J. et al. SNP web resources and their potential applications in personalized medicine. **Current drug metabolism**, v. 13, n. 7, p. 978–90, 2012.

WANG, L. L.; LI, Y.; ZHOU, S. F. A bioinformatics approach for the phenotype prediction of nonsynonymous single nucleotide polymorphisms in human cytochromes P450. **Drug Metabolism and Disposition**, v. 37, n. 5, p. 977–991, 2009.

WESTBROOK, K.; STEARNS, V. Pharmacogenomics of breast cancer therapy: an update. **Pharmacology & therapeutics**, v. 139, n. 1, p. 1–11, jul. 2013.

WORTH, C. L. et al. A structural bioinformatics approach to the analysis of nonsynonymous single nucleotide polymorphisms (nsSNPs) and their relation to disease. **Journal of bioinformatics and computational biology**, v. 5, n. 6, p. 1297–318, 2007.

WU, X. et al. Advance in Structural Bioinformatics. v. 827, p. 123–141, 2015.

ZANGER, U. M.; SCHWAB, M. Cytochrome P450 enzymes in drug metabolism: regulation of gene expression, enzyme activities, and impact of genetic variation. **Pharmacology & therapeutics**, v. 138, n. 1, p. 103–41, abr. 2013.

ZHANG, X.; KUIVENHOVEN, J. A.; GROEN, A. K. Forward Individualized Medicine from Personal Genomes to Interactomes. **Frontiers in Physiology**, v. 6, n. December, p. 1–11, 2015.

ANEXOS

Atividades desenvolvidas durante o mestrado*Artigo Científico em Colaboração*

Sensors **2015**, *15*, 15562–15577; doi:10.3390/s150715562

OPEN ACCESS

sensors

ISSN 1424-8220

www.mdpi.com/journal/sensors

Article

A Sensitive and Selective Label-Free Electrochemical DNA Biosensor for the Detection of Specific Dengue Virus Serotype 3 Sequences

Natália Oliveira ^{1,*}, Elaine Souza ², Danielly Ferreira ¹, Deborah Zanforlin ¹, Wessulla Bezerra ¹, Maria Amélia Borba ¹, Mariana Arruda ¹, Kennya Lopes ³, Gustavo Nascimento ¹, Danyelly Martins ^{1,4}, Marli Cordeiro ⁴ and José Lima-Filho ^{1,4}

¹ Laboratório de Imunopatologia Keizo Asami (LIKA), Universidade Federal de Pernambuco-UFPE, Av. Prof. Moraes Rego, s/n, Campus da UFPE, 50670-901 Recife, PE, Brazil; E-Mails: daniellylsantos@hotmail.com (D.F.); deborahzanforlin@gmail.com (D.Z.); wessullas@yahoo.com.br (W.B.); mariameliaiorbba@gmail.com (M.A.B.); mariana.s.arruda@gmail.com (M.A.); galvesn23@gmail.com (G.N.); bruneska@prospecmol.org (D.M.); joseluiz60@mac.com (J.L.-F.)

² Universidade Federal de Alagoas (UFAL), Campus Arapiraca, Av. Manoel Severino Barbosa, s/n, Bom Sucesso, 57.309-005 Arapiraca, AL, Brazil; E-Mail: elainevms@yahoo.com.br

³ Departamento de Virologia e Terapia Experimental (LAVITE), Centro de Pesquisas Aggeu Magalhães (CPqAM), Fundação Oswaldo Cruz (Fiocruz)—Pernambuco, Av. Professor Moraes Rego, s/n, Campus da UFPE, 50.670-420 Recife, PE, Brazil; E-Mail: kennya.genne@bol.com.br

⁴ Departamento de Bioquímica, Universidade Federal de Pernambuco-UFPE, Av. Professor Moraes Rego, s/n, Campus da UFPE, CEP: 50670-901 Recife, PE, Brazil; E-Mail: Marli@cpqam.fiocruz.br

* Author to whom correspondence should be addressed; E-Mail: nataliacybelle.89@gmail.com
Tel: +55-81-2101-2508; Fax: +55-81-2126-8000.

Academic Editor: Stephen Holler

Received: 1 March 2015 / Accepted: 23 June 2015 / Published: 1 July 2015

Abstract: Dengue fever is the most prevalent vector-borne disease in the world, with nearly 100 million people infected every year. Early diagnosis and identification of the pathogen are crucial steps for the treatment and for prevention of the disease, mainly in areas where the co-circulation of different serotypes is common, increasing the outcome of dengue hemorrhagic fever (DHF) and dengue shock syndrome (DSS). Due to the lack of fast and inexpensive methods available for the identification of dengue serotypes, herein we report

Resumos Apresentados em Simpósio Internacional

- OLIVEIRA, N. C. L.; SOUZA, E. V. M.; ZANFORLIN, D. M. L.; **SOUTO-MAIOR, M. A. C. B.**; ROCHA FILHO, S. L. G.; SILVA, M. M.; SOUZA JUNIOR, V. R.; NASCIMENTO, G. A.; LIMA FILHO, J. L.. Comparison of two methods to attach ssDNA onto carbon electrodes modified with poly-l-lysine. V Simpósio Internacional em Diagnóstico e Terapêutica e VIII Jornada Científica do LIKA (ago/2014)
- ZANFORLIN, D. M. L.; **SOUTO-MAIOR, M. A. C. B.**; ROCHA FILHO, S. L. G.; RODRIGUEZ, B. A. G.; OLIVEIRA, N. C. L.; FERREIRA, D. S. C.; NASCIMENTO, G. A.; BRUNESKA, D.; AGUIAR, J. A. O.; LIMA FILHO, J. L.. Biosensor based in polyaniline nanofibers, silver nanoparticles and graphene oxide, for HPV detection in Real Samples. V Simpósio Internacional em Diagnóstico e Terapêutica e VIII Jornada Científica do LIKA (ago/2014)
- **SOUTO-MAIOR, M. A. C. B.**; MELO NETO, R. P. E.; SILVA, M. M.; SOUZA JUNIOR, V. R.; OLIVEIRA, N. C. L.; ROCHA FILHO, S. L. G.; FERREIRA, D. S. C.; NASCIMENTO, G. A.; ZANFORLIN, D. M. L.; LIMA FILHO, J. L.. In silico analysis of molecular target for diagnosis of HIV by PEDOT-BASED biosensor. V Simpósio Internacional em Diagnóstico e Terapêutica e VIII Jornada Científica do LIKA (ago/2014)
- **MARIA AMÉLIA CARLOS SOUTO MAIOR BORBA**; GLAUBER MOREIRA LEITÃO; CARLOS HENRIQUE CASTELLETTI; DANYELLY BRUNESKA GONDIM MARTINS. CYP2D6 Alleles Polymorphisms And Tamoxifen Responses In Breast Cancer - A Bioinformatics Approach. VI Simpósio Internacional em Diagnóstico e Terapêutica e IX Jornada Científica do LIKA (ago/2015)
- RENATO PESSOA E MELO NETO, **MARIA AMÉLIA CARLOS SOUTO MAIOR BORBA**, CARLOS HENRIQUE CASTELLETTI, DANYELLY BRUNESKA GONDIM MARTINS, JOSÉ LUIZ DE LIMA FILHO. Template-Based Modeling As Tool For Analyzing HPV11 E2 DNA Binding Domain Secondary Structure. VI Simpósio Internacional em Diagnóstico e Terapêutica e IX Jornada Científica do LIKA (ago/2015)

Prêmios

- **3º Colocado Categoria Oral** - V Simpósio Internacional em Diagnóstico e Terapêutica e VIII Jornada Científica do LIKA (ago/2014). OLIVEIRA, N. C. L.; SOUZA, E. V. M.; ZANFORLIN, D. M. L.; **SOUTO-MAIOR, M. A. C. B.**; ROCHA FILHO, S. L. G.; SILVA, M. M.; SOUZA JUNIOR, V. R.; NASCIMENTO, G. A.; LIMA FILHO, J. L. **Comparison of two methods to attach ssDNA onto carbon electrodes modified with poly-L-lysine.**
- **1º Colocado Categoria Oral** - V Simpósio Internacional em Diagnóstico e Terapêutica e VIII Jornada Científica do LIKA (ago/2014). ZANFORLIN, D. M. L.; **SOUTO-MAIOR, M. A. C. B.**; ROCHA FILHO, S. L. G.; RODRIGUEZ, B. A. G.; OLIVEIRA, N. C. L.; FERREIRA, D. S. C.; NASCIMENTO, G. A.; BRUNESKA, D.; AGUIAR, J. A. O.; LIMA FILHO, J. L. **Biosensor based in polyaniline nanofibers, silver nanoparticles and graphene oxide, for HPV detection in Real Samples.**
- **1º Colocado Categoria Poster** - V Simpósio Internacional em Diagnóstico e Terapêutica e VIII Jornada Científica do LIKA (ago/2014) **SOUTO-MAIOR, M. A. C. B.**; MELO NETO, R. P. E.; SILVA, M. M.; SOUZA JUNIOR, V. R.; OLIVEIRA, N. C. L.; ROCHA FILHO, S. L. G.; FERREIRA, D. S. C.; NASCIMENTO, G. A.; ZANFORLIN, D. M. L.; LIMA FILHO, J. L.. **In silico analysis of molecular target for diagnosis of HIV by PEDOT-BASED biosensor**
- **3º Colocado Categoria Oral** - VI Simpósio Internacional em Diagnóstico e Terapêutica e IX Jornada Científica do LIKA (ago/2015). **MARIA AMÉLIA CARLOS SOUTO MAIOR BORBA**; GLAUBER MOREIRA LEITÃO; CARLOS HENRIQUE CASTELLETTI; DANYELLY BRUNESKA GONDIM MARTINS. **CYP2D6 Alleles Polymorphisms And Tamoxifen Responses In Breast Cancer - A Bioinformatics Approach.**
- **1º Colocado Categoria Poster** - VI Simpósio Internacional em Diagnóstico e Terapêutica e IX Jornada Científica do LIKA (ago/2015). RENATO PESSOA E MELO NETO, **MARIA AMÉLIA CARLOS SOUTO MAIOR BORBA**, CARLOS HENRIQUE CASTELLETTI, DANYELLY BRUNESKA GONDIM MARTINS, JOSÉ LUIZ DE LIMA FILHO. **Template-Based Modeling As Tool For Analyzing HPV11 E2 DNA Binding Domain Secondary Structure.**

- **Silver Medal**, iGEM International Genetically Engineered Machine Competition, MIT (out/2014). ZANFORLIN, D. M. L.; SILVA, M. M.; GODONE, R.; **SOUTO-MAIOR, M. A. C. B.**; MELO NETO, R. P. E.; CATENA, A. S.; CARDOSO, M. F.; VERDE, F. V.; BRITO, R.; SANTOS, L.; ALBUQUERQUE, J.; LIMA FILHO, J. L.. **Colialert For Breastbotsensor**.

Participação Em Comissões

- Comissão Científica (Avaliação de Trabalhos – Poster) na XXI Semana de Biomedicina – UFPE “Biotecnologia e inovação no desenvolvimento das ciências biomédicas”. 22 a 26 de setembro de 2014. Recife-PE
- Comissão Científica (Avaliação de Trabalhos – Poster) na XXII Semana de Biomedicina – UFPE “Perspectivas no avanço científico e diagnóstico” 03 a 06 de novembro de 2015. Recife-PE.
- Comissão organizadora do V Simpósio Internacional em Diagnóstico e Terapêutica e VIII Jornada Científica do LIKA – agosto/2014
- Comissão organizadora do VI Simpósio Internacional em Diagnóstico e Terapêutica e IX Jornada Científica do LIKA – agosto/2015
- Comissão Organizadora do I Translational Breast Cancer International Symposium (TBCIS) – novembro/2014

Participação Em Eventos

- Participou como palestrante do I Translational Breast Cancer International Symposium (TBCIS) – 2014, ministrando a palestra intitulada: Biosensors for Breast Cancer.

- Participou como palestrante da XXII Semana de Biomedicina da UFPE “Perspectivas no avanço científico e diagnóstico” ministrando a palestra intitulada: Farmacogenômica no Câncer.

Outras Atividades

Revisão De Periódico

International Archives of Nursing and Health Care, manuscrito ID: **WZ29205792_2015-12-01.**

Estágio De Docência

Realizou estágio de docência na disciplina Bioquímica II para a graduação em Biomedicina no período de 2015.2, sob a supervisão da professora Maria da Paz Carvalho e da professora Danyelly Bruneska Gondim Martins.

Treinamento Ou Capacitação

Participou do treinamento de Configuração, Instalação e Validação do equipamento POTENCIOSTATO/GALVANOSTATO PGSTAT128N com módulo MUX (Metrohm Pensalab), com duração de 16h. 06 a 08 de julho de 2015, Recife - PE