



**Universidade Federal do Estado do Rio de Janeiro
Centro de Ciências Biológicas e da Saúde
Instituto Biomédico
Programa de Pós-Graduação em Biologia Molecular e
Celular**

**Predição computacional do efeito das mutações P703L e
F2717L na proteína HTT humana**

Silvio Caetano Alves Junior

**Orientadora: Prof^a. Dr^a. Joelma Freire de Mesquita
Coorientadora: Mestre Aloma Nogueira Rebello da Silva**

**Rio de Janeiro
2020**

Silvio Caetano Alves Junior

**Predição computacional do efeito das mutações P703L e
F2717L na proteína HTT humana**

**Orientadora: Prof^a. Dr^a. Joelma Freire de Mesquita
Coorientadora: Mestre Aloma Nogueira Rebello da Silva**

Dissertação de mestrado apresentada ao
Programa de Pós-Graduação em Biologia
Molecular e Celular da Universidade Federal
do Estado do Rio de Janeiro como parte dos
requisitos necessários à obtenção do grau de
Mestre em Biologia Molecular e Celular.

**Rio de Janeiro
2020**

Silvio Caetano Alves Junior

**Predição computacional do efeito das mutações P703L e F2717L na
proteína HTT humana**

Dissertação de mestrado apresentada ao
Programa de Pós-Graduação em Biologia
Molecular e Celular da Universidade Federal
do Estado do Rio de Janeiro como parte dos
requisitos necessários à obtenção do grau de
Mestre em Biologia Molecular e Celular.

Banca examinadora:

Prof. Dr. Cassiano Felipe Gonçalves de Albuquerque - UNIRIO
Prof. Dr. Carlos Fernando Araujo Lima de Oliveira - CESVA/FAA
Prof. Dr. Rodrigo Grazinoli Garrido - UFRJ
Prof^a. Dr^a. Kenia Balbi El-Jaick - UNIRIO
Prof. Dr. Pablo Trindade - UNIRIO

Alves Junior, Silvio Caetano.

Cutter Predição computacional do efeito das mutações P703L e F2717L na
proteína HTT humana / Silvio Caetano Alves Junior. – 2020.
45 f. : il

Dissertação (Mestrado em Biologia celular e molecular) – Universidade
Federal do Estado do Rio de Janeiro, Rio de Janeiro, 2020.

Orientadora: Profª. Drª. Joelma Freire de Mesquita.

Coorientadora: Mestre Aloma Nogueira Rebello da Silva.

Bibliografia

1.Huntingtina. 2.Síndrome de Rett. 3.In sílico. I. Predição computacional
do efeito das mutações. II. Universidade Federal do Estado do Rio de Janeiro.

CDD

Resumo

A Huntingtina é uma proteína existente nos neurônios, que transita constantemente entre o núcleo e o citoplasma, apresenta 3142 resíduos de aminoácidos em sua cadeia. Mutações na sequência de aminoácidos da Huntingtina (HTT) pode causar algumas doenças, dentre elas a síndrome de Rett. Pouco se sabe sobre as informações estruturais da HTT, devido a complexidade de se obter um método que permita sua descrição detalhada. A síndrome de Rett é caracterizada pela interrupção do desenvolvimento neuromotor, ocorrido de maneira prevalente em meninas. A causa da Síndrome de Rett se dá por uma mutação no gene que codifica a proteína-2 de ligação metil-CpG - MECP2, a variante congênita da síndrome de Rett, é causada por uma mutação no gene FOXP1. Com tudo o estudo se baseia em mutações que geram as mesmas características fenotípicas que definem a síndrome de Rett, porém em regiões diferentes à MECP2 e FOXP1, estando vinculado a dois casos da síndrome ocorrido tanto pela mutação P703L(LOPES et al., 2016) como pela F2717L(RODAN et al., 2016), apresentando resultados diferentes em cada um dos casos e mesmo assim estando dentro do quadro clínico da síndrome de Rett. O trabalho computacional nos permite realizar manipulações de menor custo, facilitando a análise relacionada a um genoma por completo, trazendo melhor clareza dos organismos e da complexidade de seus fatores biológicos. Nesse estudo objetivamos analisar as alterações funcionais e estruturais das mutações P703L e F2717L da proteína HTT por modelagem estrutural *in silico*. por meio da utilização de algoritmos e software realizamos de uma forma geral: alinhamento, identificação de padrões de sequências e modelagem da proteína e suas variantes.

ABSTRACT

Huntingtin is an existing protein in the neurons, that transits constantly between nucleus and cytoplasm, presents 3142 amino acid residues in its chain. Mutations in the HTT amino acid sequence can cause some diseases, including Rett syndrome. Little is known about the structural information of HTT, due to the complexity of obtaining a method that allows its detailed description. Rett syndrome is characterized by the interruption of neuromotor development, which is prevalent in girls. The cause of Rett's syndrome is due to a mutation in the gene that encodes the methyl-CpG-binding protein-2 - MECP2, the congenital variant of Rett's syndrome, is caused by a mutation in the FOXP1 gene. However, the study is based on studies of mutations that generate the same phenotypic characteristics that define Rett syndrome, but in different regions to MECP2 and FOXP1, being linked to two cases of the syndrome that occurred both by the P703L mutation (LOPES et al., 2016) as well as by F2717L (RODAN et al., 2016) presenting different results in each case and still being within the clinical picture of Rett syndrome. The computational work allows us to perform manipulations at a lower cost, facilitating the analysis related to a complete complete genome, bringing better clarity of the organisms and the complexity of their biological factors. In this study we aim to analyze the functional and structural changes of the HTT protein P703L and F2717L mutations by structural modeling in silico using algorithms and software in general: alignment, identification of sequential patterns and modeling of the protein and its variants.

Sumário

Introdução	Pág.9
Fisiopatologia da síndrome de Rett	Pág.10
Síndrome de Rett no sexo masculino	Pág.11
Crítérios de diagnóstico	Pág. 12
Tabela 1. Características clínicas para diagnóstico da síndrome de Rett clássica.	Pág. 12
Tabela 2. Características de suporte ao diagnóstico da síndrome de Rett.	Pág. 13
Tabela 3. Características clínicas para diagnóstico da síndrome de Rett atípica.	Pág. 14
Doença de Huntington	Pág. 15
Tabela 4 . Características para diagnóstico da doença de Huntington.	Pág. 15
Huntingtina.	Pág. 16
In Sílico	Pág.18
Algoritmo e Modelagem	Pág. 18
Casos clínicos das variantes	Pág. 20
Objetivo geral	Pág. 21
Objetivos específicos	Pág. 21

Metodologia	Pág. 21
Compilação da Huntingtina e suas mutantes.	Pág. 22
Predição funcional das mutações na proteína Huntingtina.	Pág. 22
Modelagem computacional comparativa e ab initio da estrutura tridimensional das mutações da Huntingtina	Pág. 22
Alinhamento	Pág. 23
Padrões e domínios	Pág. 23
Padrões de conservação	Pág. 23
Mutantes	Pág. 24
Resultado e Discussão	Pág. 24
Predições de 13 Algoritmos Sobre o Efeito da substituição de aminoácidos.	Pág. 25
Tabela 1.	Pág. 26
Tabela 2.	Pág. 26
Modelagem computacional.	Pág. 27
Figura 1.	Pág. 28
Predição de regiões de conservação.	Pág. 29
Figura 2.	Pág. 30
Conclusão.	Pág. 31
Bibliografia.	Pág. 33

Introdução

A Huntingtina é uma proteína amplamente expressa em neurônios. Esta proteína de 3142 resíduos de aminoácidos, transita constantemente entre o núcleo e o citoplasma e desempenha um papel essencial no desenvolvimento humano. Consequentemente, alterações nesta proteína podem levar ao surgimento de patologias graves, dentre as quais, a mais conhecida é a doença de Huntington, a qual é causada pelo excesso de trinucleotídeos consecutivos CAG que codificam uma região de poliglutamina N-terminal na sequência genética da Huntingtina (Futter et al., 2009). Entretanto, estudos demonstram que diferentes mutações que podem ocorrer na sequência de aminoácidos da Huntingtina, provocando outras patologias, como por exemplo, a síndrome de Rett, causada tanto pela mutação P703L(LOPES et al., 2016) como pela F2717L(RODAN et al., 2016), que resultam em características clínicas distintas.

O Brasil ainda é um país carente em dados epidemiológicos sobre a síndrome de Rett. Contudo, estudos anteriores relataram o quadro clínico de pessoas portadoras desta síndrome, a exemplo do primeiro caso identificado por Rosemberg et al., 1986. A incidência da síndrome de Rett é estimada entre 1:10.000 e 1:15.000 meninas com casos da doença, admitindo ser uma das causas mais observadas de deficiência mental severa que afeta o sexo feminino (Schwartzman JS., 2003). Um estudo mais recente realizou o sequenciamento de 80 pacientes com síndrome de Rett, encontrando vinte e nove diferentes variações em 49 pacientes. Destes, 34 (69,4%) pacientes apresentaram a forma clássica da síndrome, enquanto 15 (30,6%) apresentaram formas atípicas da doença. Além disso, 13 das 29 mutações detectadas representam novas variações de sequência (Leila Schuindt et al., 2010).

A ampla distribuição da em diferentes organelas e a alta capacidade de interações moleculares da Huntingtina, no entanto, dificulta o esclarecimento das suas funções biológicas, bem como dos mecanismos moleculares que levam ao desenvolvimento das patologias associadas (Matthias Michalek et al.,

2013). Portanto, a realização de estudos voltados à compreensão das funções desta proteína, bem como dos mecanismos fisiopatológicos associados à síndrome de Rett possuem uma relevância clínica significativa e poderão contribuir para o desenvolvimento de novas estratégias terapêuticas.

Fisiopatologia da Síndrome de Rett

A síndrome de Rett é definida como um grupo de sinais e sintomas de manifestações clínicas, independentemente da etiologia, que caracteriza a suspensão do desenvolvimento neuromotor, prevalente em crianças do sexo feminino. Aparentemente as manifestações clínicas são perceptíveis a partir de 6 a 8 meses de vida. De forma geral, os primeiros sinais clínicos são os movimentos estereotipados com as mãos e a microcefalia, com posterior desenvolvimento de convulsões e retardo mental. Embora seja raro, há relatos de meninos afetados, os quais apresentaram classicamente mosaicismos somáticos ou um cromossomo X extra (Moog et al., 2003).

A forma clássica da Síndrome de Rett é decorrente de uma mutação no gene que codifica a proteína-2 de ligação metil-CpG - MECP2; no cromossomo Xq28. O MECP2 é uma proteína associada à cromatina que se liga a sequências CpGs metilados, podendo ativar e reprimir a transcrição. A principal função de MECP2 é condensar a cromatina para montagem de novas estruturas secundárias, sendo necessária para a maturação e desenvolvimento dos neurônios (Swanberg et al., 2009).

A variante congênita da síndrome de Rett é causada por uma mutação no gene FOXP1 localizado no cromossomo 14q13. O gene FOXP1 codifica um fator de transcrição do desenvolvimento com atividades repressoras (Murphy et al., 1994 ; Li et al., 1995), causando um distúrbio do desenvolvimento neurológico grave com características da síndrome de Rett clássica, mas com início mais precoce nos primeiros meses de vida, diferindo, neste aspecto, do desenvolvimento tardio visto na forma clássica. Não há cura para esta condição, embora medicamentos, fisioterapia, psicoterapia e suporte nutricional ajudem a controlar os sintomas, evitar complicações e melhorar a qualidade de vida (Mitter et al., 2018)

Por outro lado, a síndrome de Rett definida como forma atípica, engloba os demais casos que se enquadram nas características clínicas da síndrome, embora nem todas as manifestações clínicas estejam necessariamente presentes. Além disso, as manifestações podem se dar em idade diferente da esperada, e não envolver a mutação do gene MECP2 nem de sua variante congênita FOXP1 (Zahraa, et al., 2020).

Síndrome de Rett no sexo masculino

Os primeiros casos de Síndrome de Rett em meninos foram descritos em 1990, um deles por Coleman e outros dois por Philippart. De um modo geral, meninos com síndrome de Rett sofrem de morte prematura ou, então, aborto espontâneo durante a gestação, devido ao comprometimento grave do desenvolvimento neurológico (Topçu et al., 2002). No entanto, é descrito que meninos podem desenvolver a síndrome em três diferentes situações: 1) meninos que não apresentam alterações em seu cariótipo para o sexo masculino (46, XY), mas possuem mutações no gene MECP2. Neste caso, pode ocorrer duplicação do gene MECP2 ou ainda, ausência de cópia do gene para contrabalançar a mutação do alelo, o que leva a manifestação de sintomas severos nos pacientes (COSTA, Marcella., 2016); 2) meninos que apresentam o cariótipo (47, XY), caracterizado pela presença da Síndrome de Klinefelter, o que está associado à mutação do alelo do gene MECP2 em um dos cromossomos X, formando um quadro clínico de duas síndromes concomitantemente, a de Klinefelter com a Síndrome de Rett (SCHWARTZMAN JS et al., 1998); 3) meninos que apresentam um cariótipo normal (46, XY) nos quais apenas algumas de suas células somáticas apresentam um alelo mutado, podendo ocorrer com isso o mosaicismo somático (Topçu et al., 2002).

Meninos com Síndrome de Rett apresentam pouca chance de sobreviver, atingindo no máximo dois anos de idade. Esse aspecto difere significativamente do que é observado nas meninas, cuja expectativa de vida varia entre 60 e 70 anos. Esta diferença notória se dá pela presença de dois cromossomos X nas meninas (Hagberg, et al., 2002).

Critérios de diagnóstico

O diagnóstico da Síndrome de Rett, especialmente na sua forma clássica, é realizado a partir das características clínicas. Exames para detectar alterações genéticas no gene para MECP2 ou de sua variante FOXP1 são realizados apenas para diagnóstico da síndrome na forma atípica. O acordo internacional aceito e formado nos últimos anos foi publicado por Hagberg et al. em 2002, e é o que tem sido utilizado como modelo padrão no diagnóstico da síndrome de Rett, conforme mostrado nas tabelas 1 e 2.

Tabela 1. Características clínicas para diagnóstico da síndrome de Rett clássica.

Norma para diagnóstico da síndrome de Rett Clássica.

I – Condições:

- 1 - Progressão pré-natal e perinatal supostamente normal;
- 2 - Progressão psicomotora supostamente normal no período dos cinco primeiros meses após o nascimento;
- 3 - Ausência inicial da característica de microcefalia.

II – Surgimento posterior ao período normal do desenvolvimento:

- 1 - Comprometimento do desenvolvimento craniano no período entre o 5º e o 48º mês de vida;
 - 2 - Comprometimento das habilidades manuais adquiridas no período do 5º à 30º mês de vida, com o surgimento subsequente de movimentos estereotipados das mãos (por ex. gestos como torcer ou lavar as mãos);
 - 3 - Desinteresse precoce do ambiente de âmbito social;
 - 4 - Marcha e movimentos do tronco desordenados ou diminuição do tônus muscular;
 - 5 - Comprometimento severo das linguagens seja ela para se expressar ou de receptividade, com severo retardo psicomotor.
-

Fonte: The Rett Syndrome Diagnostic Criteria Work Group (1988) & DSM-IV-TR (2002)

Tabela 2. Características de suporte ao diagnóstico da síndrome de Rett.

Suporte não obrigatório ao diagnóstico.

- 1 - Hiperventilação;**
 - 2 - Escoliose progressiva;**
 - 3 - Crises epiléticas;**
 - 4 - Alterações diversas no eletroencefalograma de sono (EEG de sono);**
 - 5 - Distúrbios vasomotores e tróficos de extremidades (mãos e pés pequenos e frios);**
 - 6 - Retardo pondero-estatural (baixa estatura e peso subnormal);**
 - 7 - Ranger de dentes (bruxismo);**
 - 8 - Padrões de sono anormais (despertar noturno, sonolência diurna)**
 - 9 - Constipação ou prisão de ventre crônica;**
 - 10 - Distensão do abdome por ingestão de ar**
 - 11 - Sinais espásticos, rídez e distonias.**
-

Fonte: The Rett Syndrome Diagnostic Criteria Work Group (1988) & DSM-IV-TR (2002)

Nem todos os casos da síndrome de Rett necessariamente apresentam mutação em MECP2, como também existem casos de pacientes com a mutação em MECP2 que não desenvolveram características que se enquadram no diagnóstico da síndrome de Rett. Sendo assim, a característica de maior relevância para diagnóstico é um período de regressão do desenvolvimento, seguida de recuperação ou estabilização limitada. Embora a microcefalia seja um fator que dê suporte ao diagnóstico clássico, ela já não é mais vista como necessária para determinar o diagnóstico (Neul et al., 2010).

Para o diagnóstico das formas atípicas da síndrome de Rett é necessário apresentar pelo menos 3 requisitos das 6 condições de maiores relevâncias ou ainda, 5 requisitos das 11 condições de menor relevância. Pode incluir a manifestação de epilepsia precocemente, anormalidade no desenvolvimento, possível preservação da fala ou algum tipo de linguagem funcional,

manifestação tardia dos sinais e sintomas (Hagberg B, et al., 2002), conforme resumido na tabela 3.

Tabela 3. Características clínicas para diagnóstico da síndrome de Rett atípica.

Norma para diagnóstico da síndrome de Rett atípica.

I – Condições de maior relevância

- 1 - Ausência ou redução das habilidades manuais práxicas
- 2 - Redução ou perda da fala em balbucios
- 3 - Redução ou perda das habilidades comunicativas
- 4 - Desaceleração do crescimento cefálico nos primeiros anos de vida
- 5 - Padrão monótono de estereotípias manuais
- 6 - Perfil da desordem Rett: Estágio de regressão seguido de recuperação da interação social contrastando com regressão neuromotora lenta

II – Condições de menor relevância

- 1 - Irregularidades respiratórias
- 2 - Distensão abdominal / Aerofagia
- 3 - Bruxismo
- 4 - Locomoção anormal
- 5 - Escoliose / cifose
- 6 - Amiotrofia dos membros inferiores
- 7 - Contato visual intenso
- 8 - Resposta diminuída para dor
- 9 - Episódios de risadas / gritos
- 10 - Hipotermia, pés arroxeados, crescimento (usualmente) comprometido
- 11 - Distúrbios de sono, incluindo episódios de terrores noturnos

Fonte: The Rett Syndrome Diagnostic Criteria Work Group (1988) & DSM-IV-TR (2002)

Doença de Huntington

A doença de Huntington está bastante vinculada a mutação da proteína huntingtina, com isso, esse tópico vem para ressaltar características da doença de Huntington que a diferem da síndrome de Rett.

Huntington é uma doença hereditária provocada pela morte de neurônios. A consequência disso gera perda de coordenação motora, locomoção instável, perda da fala e degeneração mental até o estado de demência (Adam MP, et al., 2018). Na grande maioria das vezes os sintomas surgem entre os 30 e 50 anos de idade, porém existem relatos que a cada nova geração mais cedo ocorre a manifestação. A doença abrange entre 4 e 15 em cada 100 000 pessoas de ascendência europeia, afetando de igual forma mulheres e homens (Dayalu P, et al., 2015).

Huntington pode ser herdado tanto da mãe quanto do pai, por uma mutação genética autossômica dominante do gene que formará a proteína Huntingtina. A expansão anormal de um trato de poliglutamina na porção N-terminal da Huntingtina é a causa da doença de Huntington, devido a uma repetição triplete CAG que se encontra no gene, resulta nessa forma mutante da proteína que leva a morte dos neurônios, porém seus mecanismos não são totalmente compreendidos (Durr A, et al., 2012).

As características clínicas da doença de Huntigton são:

Tabela 4 . Características para diagnóstico da doença de Huntington.

Sintomatologia.

- 1 - Coreia (movimentos involuntários, rápidos, irregulares e sem finalidade dos membros, da face e/ou do tronco, geralmente associados à hipotonia e à diminuição da força muscular);**
 - 2 - Perda progressiva de memória;**
 - 3 - Transtornos psiquiátricos: depressão, ansiedade, irritabilidade, hipersexualidade, egocentrismo.**
 - 4 - Disartria (distúrbios da fala, mastigação e deglutição);**
-

5 - Fala incompreensível, hesitante, explosiva e desorganizada;**6 - Perda da visão periférica.**

Fonte: Duijn E, et al., 2007 - J Neuropsychiatry Clin Neurosci.

A doença de Huntington não é a causa da morte dos portadores, e sim os distúrbios originados dos sintomas em decorrência do desenvolvimento da doença, sendo progressivo o comprometimento do organismo e letal (Duijn E, et al., 2007).

Um dos motivos que a doença de Huntington e a síndrome de Rett parecem compartilhar algumas características está na possível variação que ocorre no processo de formação da Huntingtina, sendo essas variações à possível causa da doença. Já que a Huntingtina é uma proteína que interage diretamente com MECP2 e a redução da expressão de Huntingtina em todo o cérebro, leva a um defeito no transporte de BDNF, ocasionando alterações em processos transcricionais, atrofia neuropatológicas do estriado e apresentação clínica de distúrbio do movimento compulsivo mais disfunção cognitiva (McFarland, et al., 2014).

Huntingtina

Os dados sobre informações estruturais da Huntingtina ainda se encontram incompletos, devido à complexidade de se obter um método que permita sua descrição detalhada. As informações já descritas é que maior parte de sua formação é α -helicoidal com três domínios, existe nos domínios N-terminal e C-terminal múltiplas repetições do fator de alongamento 3, proteína fosfatase 2A e lipídios quinase TOR (HEAT), estes estão organizados de maneira solenoide, e se apresentam conectados por uma ponte com domínio menor contendo variadas formas de repetições em tandem (Guo, et al., 2018).

A importância de se obter mais informações da estrutura da Huntingtina em alta resolução se dá, para melhor esclarecimento da larga capacidade de diferentes interações que a Huntingtina pode realizar, para o entendimento do desenvolvimento da patogênese em relação a sua função biológica e de seus mecanismos moleculares. Aumentando ainda mais o grau de importância e de

dificuldade, ao saber de sua vasta localização em um grande número de organelas e estruturas celulares, como por exemplo, núcleo, membrana plasmática, Retículo endoplasmático, aparelho de Golgi, vesículas endocíticas e mitocôndrias. (Matthias Michalek, et al., 2013)

Estudos moleculares, demonstram que a Huntingtina é uma proteína grande de 348 KDa primordial no desenvolvimento embrionário, uma proteína nuclear altamente expressa que se liga a diversos fatores de transcrição, desenvolvendo um papel regulatório desses fatores. Apresenta um importante papel em diversas atividades da célula, como por exemplo: Autofagia, endocitose e transcrição (Guo, et al., 2018).

Sua grande capacidade de interações e suas funções biológicas sugerem que a Huntingtina em coparticipação com outras moléculas, funciona semelhante a um grande centro de interação entre proteína – proteína, sendo assim, vias intrínsecas de extrema importância para atividade do metabolismo celular pode ser comprometida devido a sua mutação (Futter, et al., 2009).

Uma das interações que a Huntingtina pode realizar, é a atividade de regular positivamente a transcrição do fator neurotrófico de BDNF, fator importante para a sobrevivência de neurônios, atividade esta de extrema importância que acaba sendo comprometida a partir do momento em que a proteína Huntingtina sofre mutação, acarretando no declínio da síntese de BDNF cortical e uma insuficiência no suporte de produção neurotrófico, levando a morte desses neurônios (Zuccato, et al., 2001).

A Huntingtina pode funcionar no interior do núcleo como uma espécie de montagem de complexos proteicos, estando voltada para processos de repressão transcricional e de processamento de RNA. Sendo assim, a proteólise da Huntingtina mutante pode causar perturbações nas funções nucleares e de suas interações do complexo proteico (Kegel, et al., 2002).

O que geralmente acontece em relação a Huntingtina mutante é o seu acúmulo no núcleo, por meio da baixa interação da porção N-terminal que é transportada do núcleo para o citoplasma, esses fragmentos por interação molecular com a região promotora de proteínas que se localiza no poro do núcleo não conseguem interagir de forma eficaz, não completando o seu transporte para o citoplasma, causando então, o acúmulo da Huntingtina no núcleo da célula (Cornett, et al., 2005).

In sílico

O trabalho computacional nos permite realizar manipulações de menor custo, facilitando a análise relacionada a um genoma por completo. Com isso, o processo se torna mais rápido para estudos diretamente direcionados a indivíduos ou populações de indivíduos, trazendo melhor clareza dos organismos e da complexidade de seus fatores biológicos. Permite o possível entendimento da história evolutiva de um conjunto de organismos ou até mesmo, traçar vias de interações entre moléculas de um determinado organismo, tecido ou tipo celular, de uma forma geral: alinhamento, identificação de padrões sequências, filogenia, genomas e biologia de sistemas. (Hugo Verli. – 1ª ED. 2014, Pag.7)

Algoritmo e Modelagem

A utilização de software e algoritmos contribui-se muito para estudos que envolvem a estrutura das biomoléculas, servindo para análise evolutiva de conservação de resíduos e estabilidade da estrutura proteica, possíveis comparações de estruturas em terceira dimensão, e na realização de seu alinhamento. Fornecendo então, melhores entendimentos sobre seu significado de conservação seja de grandes resíduos a pequenos conjuntos de sequências. (Hugo Verli. – 1ª ED. 2014, Pag.7)

Dos primordiais avanços conhecidos na pesquisa que envolve a seleção de recursos em um conjunto de aplicações de bioinformática, dois problemas principais merecem atenção, como problemas comuns no domínio do uso da bioinformática: a grande dimensionalidade da entrada e os pequenos tamanhos de amostra. De modo a solucionar esses problemas, diversas técnicas foram projetadas por pesquisadores em bioinformática, como aprendizado de máquina e mineração de dados.

De uma forma geral, muitos pesquisadores da área ainda se prendem a abordagens univariadas. A proposta do uso de algoritmos de seleção multivariada pode ser considerada uma das linhas de trabalho mais

promissoras para a comunidade de bioinformática, como por exemplo, SNPs mineração de texto e literatura e a combinação de fontes de dados heterogêneos, a metodologia se torna essencial no tratamento do caráter de alta dimensão dessas aplicações. (Yvan Saeys et al., 2007)

Os nsSNPs têm sido utilizados no estudo de uma vasta variedade de doenças. Avaliando a função das proteínas, de suas alterações em relação as regiões de ligação ao DNA e do fator de transcrição, que por sua vez, reduz a solubilidade das proteínas e desestabiliza suas estruturas. Em vista disso, compreender as consequências de alterações funcionais não-sinônimas e prever as possíveis causas e a base molecular de doenças requer a integração de o máximo de informações, sendo elas de diversas fontes heterogêneas, envolvendo sequência, dados estruturais e relações entre vias entre proteínas. Na contemporaneidade, é possível observar inúmeros estudos e recursos que permitem explorar os efeitos dos nsSNPs na estrutura terciária de proteínas e em sua funcionalidade, sendo esses recursos lançados para uso público (Alper et al., 2007).

A bioinformática sem duvida possui um papel fundamental na indústria farmacêutica, bem como em outras áreas da pesquisa científica atualmente. Na terapia medicamentosa, é importante prever interações entre as moléculas a fim de avaliar possíveis efeitos colaterais ou melhorá-las para que tenham uma melhor compatibilidade com as moléculas desejadas. Entender as mutações que podem gerar alterações na sequência de aminoácidos é de vital importância para o desenvolvimento de novos fármacos, mais eficazes no tratamento de doenças genéticas.

O projeto desenvolvido tem como foco central a utilização de ferramentas da Bioinformática para análise in silico das mutações e modelagem da Huntingtina Humana, apoiando a ideia de desenvolvimento sustentável da ciência com baixo custo e impactos ao meio, dessa forma os resultados obtidos e no projeto tem uma grande importância, para novas caracterizações de estudos e vantagens sobre possíveis fármacos, que posteriormente possam auxiliar a minimizar algumas características causadas pela mutação.

Caso clínico das variantes

O trabalho se iniciou com a análise de dois casos clínicos distintos mais que geram as mesmas características fenotípicas que definem a síndrome de Rett. Embora seja um caso clínico confirmado de síndrome de Rett, as características se enquadram na forma atípica da síndrome, em que a mutação presente na Huntingtina não envolve os genes MECP2 e FOXP1 já anteriormente mencionado. A variante P703L foi descrita como uma variante heterozigota que no caso foi de origem materna, (LOPES et al., 2016) e da mesma forma foi descrito para a variante F2717L, porém citado como uma região que deduz ser conservada (RODAN et al., 2016).

Da paciente diagnosticada com a síndrome de Rett, e que apresenta a variante P703L, foi descrito regressão do desenvolvimento por volta dos 6 meses, após o período de 2 meses a falta de interação com o ambiente, com os exames neurológicos foi descrito uma dificuldade de deglutição, distonia (contração que afeta de modo regular um ou vários músculos ou membros e, por vezes, o eixo corporal), bradicinesia sem rigidez (lentidão anormal dos movimentos voluntários), estereotipias manuais (ações repetitivas com as mãos), e com a ressonância foi relatado atrofia significativa do estriado, do córtex e vermis cerebelar, o autor relata atender os critérios do diagnóstico para síndrome de Rett, e mais oito critérios suporte (LOPES et al., 2016).

Em relação a paciente que apresenta a variante F2717L, foi descrito um atraso no seu desenvolvimento desde o nascimento, hipotonia central (redução do tônus postural), dificuldades de deglutição, bruxismo, miopia grave, hipertonia apendicular progressiva com postura distônica intermitente, estereotipias na linha média, incluindo torção nas mãos. Ela nunca teve convulsões. É importante mencionar que nesse estudo, a menina tem dois irmãos que não são diagnosticados com a síndrome, sendo observado em um deles, algumas características semelhantes a da irmã, mas que não se enquadra no quadro clínico da síndrome de Rett (RODAN et al., 2016).

Objetivo Geral

Nesse estudo objetivamos analisar as alterações funcionais e estruturais das mutações P703L e F2717L da proteína HTT por modelagem estrutural in silico.

Objetivos Específicos

- 5.1. Predição funcional de SNPs, visando estudar a patogenicidade dessas mutações;
- 5.2. Obter modelos tridimensionais por algoritmos de modelagem comparativa e ab initio;
- 5.3. Alinhamento estrutural dos modelos tridimensionais mutantes e não mutantes;

Metodologia

No estudo proposto, foi coletado dados de dois casos clínicos de síndrome de Rett atípica, decorrente da variação de um aminoácido P703L(LOPES et al., 2016) e de F2717L(RODAN et al., 2016) que compõem a proteína Huntingtina para análise in Silico, que pode determinar se essas variantes podem influenciar na estrutura tridimensional, funcional ou na estabilidade da proteína. Foram criados modelos teóricos estruturais para as variantes em análise, usando modelagem comparativa por meio das coordenadas atômicas fornecidas em bancos de dados. Os resultados obtidos, posteriormente farão parte de um banco de dados que está disponível gratuitamente em <http://bioinfogroup.com/database/> e vem sendo atualizado regularmente.

Compilação das mutações na proteína Huntingtina

A compilação das mutações na proteína Huntingtina descritas na literatura e em bancos de dados de polimorfismos não sinônimos (nsSNP) humanos. Foi então, utilizados recursos de dados para mapear o desenvolvimento de nsSNPs em estruturas de proteínas, fornecendo uma visão estrutural dos efeitos dos nsSNPs em proteínas como estabilidade, funcionalidade, interações proteína-proteína e outros problemas relacionados com a estrutura proteica. Essa análise foi realizada em treze algoritmos distintos avaliando os efeitos funcionais dos polimorfismos. São esses: MAPP, PMUT, nsSNP (Bao L et al.2005), PhD-SNP (Capriotti E et al 2006), Polyphen1 e Polyphen2 (Adzhubei Ivan A et al 2010), PON-P2, Mult Pred, SIFT (Ng PC 2003), SNAP e SNAP2 (Bromberg Y et al. 2008), PredictSNP (De Baets G et al. 2012) e PROVEAN (Choi Y et al. 2012).

Predição funcional das mutações na proteína Huntingtina

Esta abordagem permite identificar se tais mutações comprometem a estabilidade ou a atividade da proteína por predições dos efeitos funcionais usando estruturas tridimensionais de proteínas como guia. Para tanto serão utilizados algoritmos como o I-Mutant e o PolyPhen (ADZHUBEI; JORDAN; SUNYAEV, 2013).

Modelagem computacional comparativa de ab initio da estrutura tridimensional das mutações da Huntingtina

Para modelagem comparativa foram utilizadas as coordenadas atômicas das estruturas tridimensionais determinadas experimentalmente por difração de raios-X ou ressonância magnética nuclear. Quando estes dados não estavam disponíveis foram utilizados modelos teóricos criados a partir dos parâmetros termodinâmicos da proteína. Esta análise objetiva correlacionar as possíveis mudanças estruturais causadas pelas mutações e a atividade biológica. Os

modelos estruturais teóricos das sequências proteicas foram criados por modelagem comparativa utilizando os algoritmos do servidor SWISS-MODEL (WATERHOUSE et al., 2018)

Alinhamento

Para investigar a similaridade estrutural foram utilizados algoritmos de alinhamento tridimensional, como o TM-align, que pode realizar comparações da sequência de estruturas proteicas, gerando um alinhamento otimizado por meio de seus resíduos, tendo como orientação as similaridades estruturais. O alinhamento estrutural gera um valor definido como TM-score, resultante de sua similaridade. A pontuação estatística do TM é (0,1), sendo 1 correspondente a uma combinação perfeita das estruturas (ZHANG; SKOLNICK, 2005).

Padrões e domínios

Para a análise de características, padrões e domínios foram utilizados os algoritmos disponíveis no ExPASy Proteomics Server (UNIPROT, 2015).

Padrões de conservação

Análise filogenética das sequências da proteína Huntingtina humana e de outros organismos visando evidenciar regiões conservadas na evolução (BEREZIN et al., 2004). O Consurf destaca-se por ser uma ferramenta capaz de projetar uma estimativa da conservação evolutiva de aminoácidos de uma proteína baseado em um banco de dados, é uma das ferramentas da bioinformática que pontua em um esquema de cores, a superfície da molécula de interesse, e essa pontuação está vinculada a relações filogenéticas da proteína e da sequência de seus homólogos. Sendo, portanto, os aminoácidos considerados de melhor relevância, aqueles em que ao longo de sua história evolutiva permaneceram conservados tanto para a estrutura, quanto para a sua função na atividade da proteína (KREBS; DE MESQUITA, 2016).

Mutantes

As estruturas mutantes foram obtidas por mutagênese in silico usando o algoritmo VMD do Linux.

Resultado e Discussão

Em concordância a casos que já foram descritos sobre a síndrome de Rett, relacionado aos efeitos da substituição dos aminoácidos das regiões P703L e F2717L. Esperava-se que todos os algoritmos fossem capazes de identificar como deletério as variantes, visto que, as regiões de interesse para o estudo já estão correlacionadas com casos clínicos da síndrome de Rett.

É de grande relevância salientar, a existência de diferentes características sobre os grupos de algoritmos utilizados na interpretação de suas previsões. De modo que dentre seus métodos de extrair informações contidas nos alinhamentos de múltiplas sequências, os algoritmos foram agrupados e diferenciados da seguinte forma: Os algoritmos que estão voltados para a previsão estrutural, funcional e de estabilidade da proteína.

Os algoritmos utilizados que estão voltados para análise estrutural são: Phd-SNP, SIFT, SNAP, Pmut, PON-P2; Os algoritmos que estão voltados não apenas para análise estrutural mais também prevê a função exercida pela proteína são: Provean, MAPP, MultiPred, Predict SNP; Os algoritmos que levam em consideração a predição voltada para a estabilidade da proteína em diferentes condições são: I-mutant, o Polyphen 1 e o Polyphen 2.

Foi então realizado predições com 13 algoritmos distintos e distribuídos em duas tabelas para melhor visualização. Na primeira tabela contendo seis algoritmos, foram agrupados os que estão relacionados à questão estrutural da proteína [Tabela 1], e a outra contendo sete algoritmos, que estão não só voltados para a questão estrutural, mas também leva em consideração a função da proteína [Tabela 2] e ainda os três últimos algoritmos da segunda tabela estão voltados para a estabilidade da proteína em diferentes condições de testes. Em ambas as tabelas foram adotadas a seguinte nomenclatura:

neutra (do inglês, Neutral), deletério (do inglês, Deleterious) e indeterminado (do inglês, Unknown). [7.1]

Predições de 13 Algoritmos Sobre o Efeito da substituição de aminoácidos.

Dos seis algoritmos utilizados para a predição da função da proteína Huntingtina em relação às variantes, cinco deles determinaram a variante P703L como deletério e uma das seis, o PON-P2 não foi capaz de realizar a predição da variante como deletério ou neutro, sendo então classificado como indeterminado. Já em relação a variante F2717L, três algoritmos foram capazes de determinar como deletério e apenas o SNAP-1 e SNAP-2 determinou a variante como neutro, e mais uma vez o algoritmo PON-P2 foi incapaz de prever como deletério ou neutro a variante F2717L, classificando como indeterminado (Tabela 1).

Na tabela 2 observamos o resultado de sete algoritmos, como dito anteriormente os quatro primeiros estão voltados para a função exercida pela proteína Huntingtina e os três últimos a estabilidade da proteína Huntingtina. Três desses algoritmos vinculados a função determinaram a variante P703L como deletério e apenas o MAPP determinou como neutro. A variante F2717L foi determinada como deletério por todos os quatro algoritmos que prevê a função da proteína. Os três últimos algoritmos vinculados com a estabilidade da proteína Huntingtina, o I-Mutant e o Polyphen-1 e 2, preveem como deletério tanto a variante P703L quanto a variante F2717L.

Tabela 1. Algoritmos Estruturais.

	PhD SNP	SFIT	SNAP-1	SNAP-2	PMUT	Pon-P2
P703L	Deletério	Deletério	Deletério	Deletério	Deletério	Indeterminado
F2717L	Deletério	Deletério	Neutro	Neutro	Deletério	Indeterminado

Tabela 2. Algoritmos Funcionais e de Estabilidade.

	Provean	MAPP	Mult Pred	Predict SNP	I-Mutant	PolyPhen-1	Polyphen-2
P703L	Deletério	Neutro	Deletério	Deletério	Deletério	Deletério	Deletério
F2717L	Deletério	Deletério	Deletério	Deletério	Deletério	Deletério	Deletério

Tabela 1 e 2. – Predições das variantes P703L / F2717L em determinação de suas estruturas, funções e estabilidade.

A importância da utilização de mais de um algoritmo para a validação e determinação do efeito das Huntingtina variantes e suas funções, está na relação dos algoritmos com suas diferentes previsões. Conforme descrito por Tugba, et al., 2016 alteração nas propriedades físico-químicas (por exemplo, polaridade, hidrofobicidade, energia de solvatação, etc.) causadas pela mutação leva a um distúrbio no equilíbrio energético na ação de empacotamento da proteína e da sua atividade exercida, podendo estabilizar ou desestabilizar a proteína. No entanto, deve-se levar em consideração o posicionamento espacial da mutação, em análise da magnitude dessa variação em relação a proporção da proteína, pra definir se houve alterações estruturais significativas que comprometam sua função.

Foi observado que alguns algoritmos determinaram as variantes como neutro. Na análise feita pelo algoritmo SNAP-1 e 2 que classificou a variante F2717L como neutra para as alterações da estrutura da Huntingtina, em

relação às características moleculares, sugerem que seja pelo fato do posicionamento espacial ou da menor variação entre as moléculas, o algoritmo não reconheceu essa variante como suficiente para gerar alterações na estrutura da proteína. Já o algoritmo PON-P2 não gerou uma determinação nas duas variantes, possivelmente pela dificuldade de se encontrar nos bancos de dados modelos completos da Huntingtina com finalidade comparativa. Na predição sobre a função da proteína, apenas o algoritmo MAPP para a variante P703L identificou como neutra, para ambas as variantes, sabe-se que estão vinculadas com casos clínicos da síndrome de Rett.

Modelagem computacional.

As estruturas tridimensionais da proteína Huntingtina nativa e suas variantes foram visualizadas pelo software PyMOL, que é um software para computador, um sistema de visualização molecular criado por Warren Lyford DeLano. O que permite produzir imagens 3D da proteína e de suas variantes. Posteriormente foi criados modelos estruturais teóricos de suas sequências proteicas, por meio de modelagem comparativa utilizando os algoritmos OS DO servidor SWISS-MODEL (WATERHOUSE ET AL., 2018). E através da análise dos resultados in silico se esperava encontrar correlações moleculares entre as alterações estruturais da Huntingtina. Possivelmente essas correlações moleculares podem auxiliar em novos estudos relacionados com o desenvolvimento de doenças motoras e interações de fármacos.

As modelagens foram formadas a partir das sequencias de seus aminoácidos, foram então formados três modelos tridimensionais, a Huntingtina natural e suas variantes [Figura 1.]. Visualmente é possível observar alterações nas estruturas das variantes em relação a natural. Por ser uma proteína grande e de difícil cristalização, mesmo os resultados de suas sequências nos bancos de dados se encontram incompletos, contendo assim alguns espaços entre alguns aminoácidos. Ainda assim a modelagem foi realizada.

Na análise comparativa dos modelos podemos notar que a variante F2717L, na porção inferior se encontra melhor conservada em sua morfologia estrutural em relação a Huntingtina natural, quando comparamos com a

morfologia estrutural na porção inferior da variante P703L, onde se encontra o primeiro domínio. O esquema de cores nos permite diferenciar e determinar os domínios da proteína Huntingtina que sofreu maior alteração, justamente onde se encontra posicionalmente a variante P703L, que será mais bem destacada na próxima figura da predição da conservação.

No modelo computacional da variante F2717L, a alteração morfológica da estrutura proteica da Huntingtina é moderada, se observamos no domínio pintado de azul, no lado direito da gravura, a existência de uma moderada alteração quando comparamos com a Huntingtina natural, o que não é observado na estrutura proteica da variante P703L.

Da mesma forma que percebemos uma diferença morfológica das estruturas entre as variantes P703L e F2717L, corroborando com os resultados descritos nas tabelas, possivelmente essas distintas alterações explique o que foi observado também com relação as diferenças clínicas entre os casos, justamente pelo fato de uma das variantes a F2717L apresentou melhor conservação estrutural de seus domínios em relação a conservação estrutural da variante P703L.

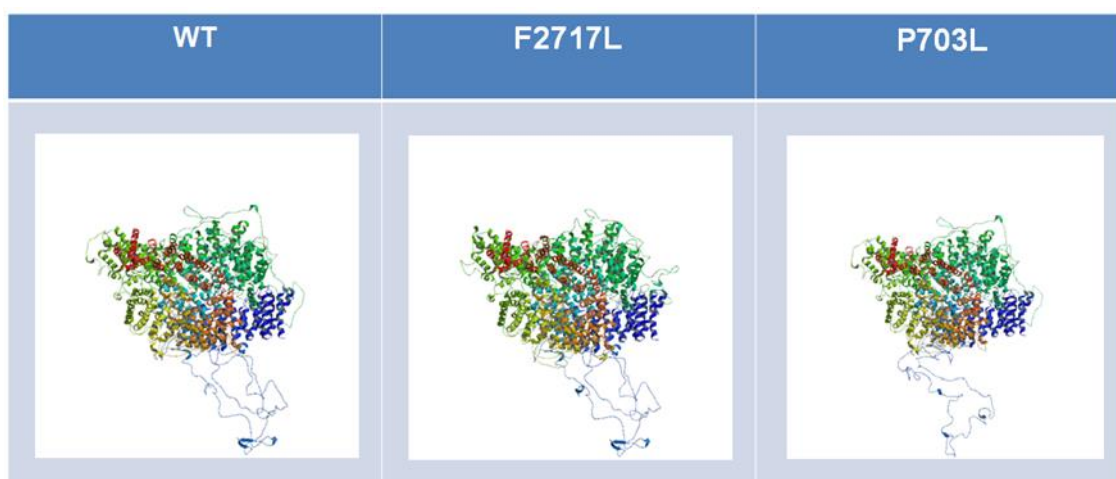


Figura 1. – Predições das variantes P703L / F2717L em determinação de suas funções.

A comparação dos resultados da tabela dos algoritmos [Tabela 1.] com o modelo, explica o motivo que os algoritmos SNAP-1 e SNAP-2 foram incapazes

de determinar como deletério a variante F2717L. A variação foi mínima em relação à Huntingtina nativa. Conforme os resultados de Guo, et al., 2018, e dos resultados gerados pelo nosso grupo nos sugerem ainda, que as alterações pela variantes P703L foram aparentemente mais comprometedoras não só em sua estrutura mais principalmente na atividade exercida pela proteína, isso por se apresentar na porção N-terminal que está de coloração azul, e ainda por se encontrar próximo do sítio de interação com o ligante, o que será melhor observado na predição de conservação.

Predição de regiões de conservação.

Para análise de conservação estrutural foi utilizado o algoritmo Consurf, capaz de determinar as regiões de conservação evolutivas e o seu grau de conservação. Permitindo então, estabelecer um nível de conservação das regiões onde se encontram as variantes, auxiliando no entendimento de como a variação da região em estudo pode influenciar tanto estruturalmente, bem como, na atividade exercida pela proteína Huntingtina.

O algoritmo Consurf classifica as regiões da seguinte forma: Azul como variável, branco como intermediária e turquesa como conservada. A imagem a seguir mostra a Huntingtina em quatro ângulos distintos para facilitar a análise e localização das variantes. Assim, foi possível destacar não só o posicionamento das variantes, mas também, mostrar onde se concentra o maior grau de conservação na proteína Huntingtina, e podemos observar que maior parte da proteína o algoritmo não foi capaz de determinar o grau de conservação na coloração amarelo, essas regiões podem ser resultado da interferência causada pela falta de informações completas nos bancos de dados. Na imagem da visão posterior pode se observar maior grau de conservação da porção mais distal a direita da Huntingtina, e do mesmo modo na imagem da visão superior da Huntingtina maior grau de conservação na porção mais inferior à direita da Huntingtina.

Foi destacada então a região referente às variantes P703L e F2717L com a seta vermelha, onde se pode observar que a variante P703L está localizada na porção, onde se encontra um maior agrupamento estrutural que

foi classificado com grau indeterminado em sua conservação, diferente do que observamos na variante F2717L, que se apresenta na porção estrutural que se obtém maior grau de variáveis diferente do que se imaginava ao descrever RODAN et al., 2016 que sugeriu ser uma região de estrutura conservada.

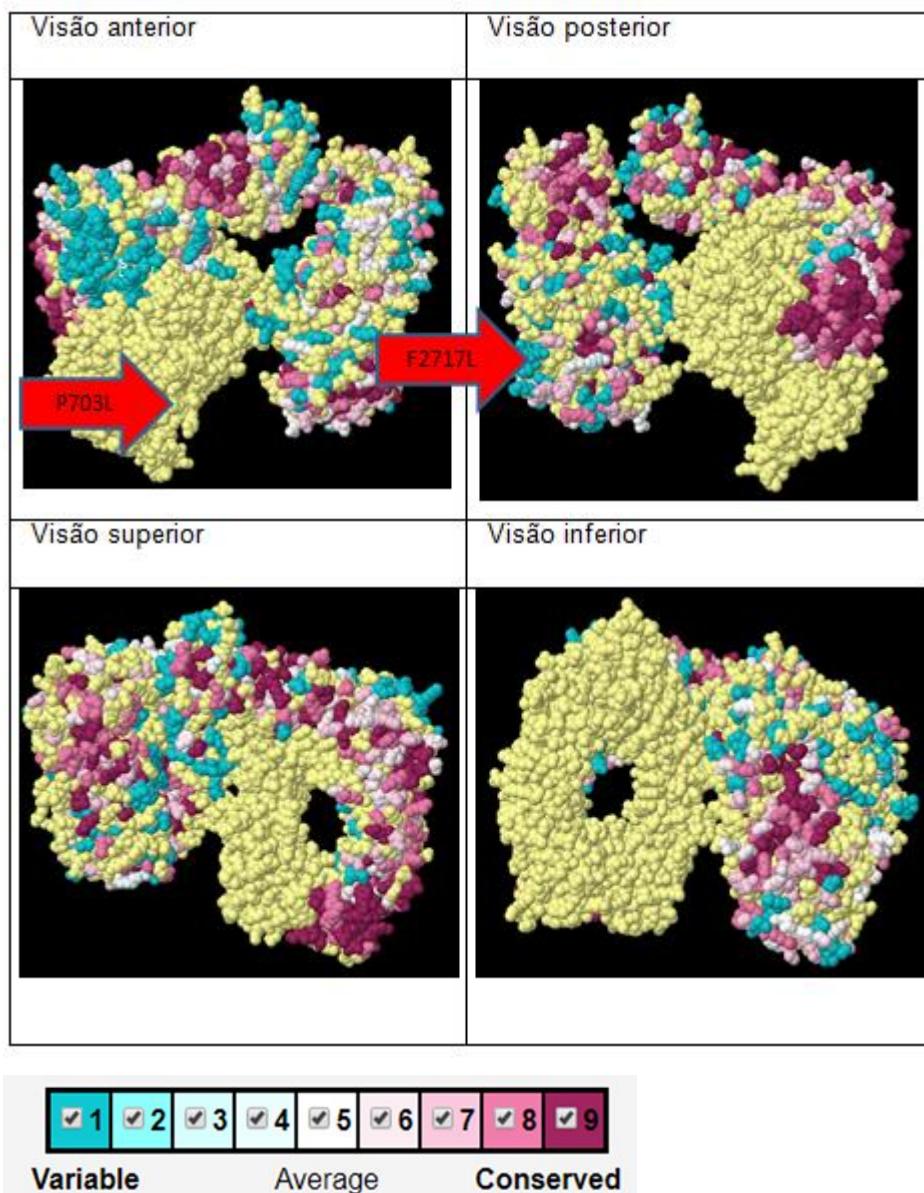


Figura 2. – Análise de conservação dos aminoácidos que formam a Huntingtina, sendo representado o grau de conservação de 1 a 9, dentro de um gradiente de cores, sendo determinado como azul variável, branco intermediária e turquesa conservada, (a cor amarela são região indeterminadas). Visto de em quatro ângulos distintos: anterior, posterior, superior e inferior. Em destaque com uma seta vermelha as regiões relacionadas as mutantes P703L e F2717L.

Esses resultados auxiliam no entendimento não só sobre a estrutura e o grau de conservação onde se encontram as variantes, mas por outro lado nos permite sugerir que a variante F2717L, por se apresentar numa porção classificada como variável, possivelmente nesse caso o comprometimento da função e estrutura da proteína Huntingtina pode ser dita como branda e em pesquisas mais detalhadas, obter melhores resultados que possam minimizar algumas características clínicas, por meio de intervenção terapêutica.

Conclusão.

Tendo em vista os relevantes aspectos mencionados, os algoritmos em conjunto, foram capazes de determinar as variantes P703L e F2717L como deletérias, corroborando com os dados já antes conhecidos que vinculam essas mutações a síndrome de Rett. Esses dados somados aos modelos tridimensionais sugerem que essas alterações estruturais da Huntingtina humana, podem sim estar vinculadas com as possíveis características que levam a causa da síndrome de Rett, por ser uma proteína de extrema importância ao sistema neurológico, e de alta afinidade de interação com outras proteínas nucleares, que a integra em diversas vias metabólicas, pode acarretar na interferência ou ainda no bloqueio dessas vias de interação levando assim as causas da Síndrome de Rett.

Com a análise das características morfológicas da proteína, foram vistas alterações causadas pelas variáveis P703L ou F2717L, mesmo sendo incapaz de realizar o alinhamento das variantes com a normal, isso se dá, pelo fato dos bancos de dados apresentarem alguns espaços na codificação de sua estrutura molecular. Possivelmente se compararmos visualmente as estruturas das variantes com a normal se pode observar sugestivamente uma moderada conservação da estrutura da F2717L em relação a P703L, o que explicaria o motivo por que na análise dos algoritmos SNAP-1 e SNAP-2, foram determinados F2717L como neutro.

A análise permitiu identificar as variantes em domínios inversos na própria proteína Huntingtina, sugerindo possíveis características distintas da

Síndrome de Rett em relação ao posicionamento das variantes, e se observarmos em detalhes o que foi descrito por LOPES et al., 2016 em relação a variante P703L e RODAN et al., 2016 da variante F2717L. As diferenças existentes leva a dedução de que a variante F2717L por ser potencialmente mais conservada, e ser definida como menos comprometedora de algumas ações metabólicas da célula nervosa em relação a P703L.

O resultado de conservação caracterizou a região da variante P703L como indeterminada e a região F2717L de fato é determinada como variável, modificando o modo de interpretação antes mencionado dentro do caso clínico como conservada, as variantes que se apresentam em regiões mais conservadas, possivelmente apresentam um papel fundamental seja para estrutura ou para função da proteína e adotando características clínicas mais comprometedoras.

Bibliografia

Adam MP, Ardinger HH, Pagon RA, Wallace SE, Bean LJH, Stephens K, Amemiya A, editors. SourceGeneReviews® [Internet]. Seattle (WA): University of Washington, Seattle; 1993-2020. **Huntington Disease**. 1998 Oct 23 updated 2018 Jul 5. Department of Medical Genetics, University of British Columbia, Vancouver, British Columbia, Canada.

Adzhubei Ivan A; Schmidt Steffen; Peshkin Leonid; Ramensky Vasily E; Gerasimova Anna; Bork Peer; Kondrashov Alexey S; Sunyaev SR. **A method and server for predicting damaging missense mutations**. Nat Methods. 2010.

ADZHUBEI, I.; JORDAN, D. M.; SUNYAEV, S. R. **Predicting functional effect of human missense mutations using PolyPhen-2**. **Current protocols in human genetics** / editorial board, Jonathan L. Haines ... [et al.], v. Chapter 7, p. Unit7.20, jan. 2013.

Alper Uzun, Chesley M. Leslin, Alexej Abyzov, and Valentin Ilyin* **Structure SNP (StSNP): a web server for mapping and modeling nsSNPs on protein structures with linkage to metabolic pathways**. Nucleic Acids Res. 2007 Jul.

Baets G, Van Durme J, Reumers J, Maurer-Stroh S, Vanhee P, Dopazo J, et al. **SNPeffect 4.0: On-line prediction of molecular and structural effects of protein-coding variants**. Nucleic Acids Res. 2012.

Bao L, Zhou M, Cui Y. **nsSNPAnalyzer: Identifying disease-associated nonsynonymous single** nucleo- 2005.

BEREZIN, C. et al. **ConSeq: the identification of functionally and structurally important residues in protein sequences**. Bioinformatics (Oxford, England), v.20, n. 8, p. 1322– 4, maio 2004.

Brazilian Journal of Psychiatry version ISSN 1809-452X Rev. Bras. Psiquiatr. vol.25 no.2 São Paulo June 2003.

Bromberg Y, Yachdav G, Rost B. **SNAP predicts effect of mutations on protein function.** Bioinformatics. 2008.

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. 2006.

Choi Y, Sims GE, Murphy S, Miller JR, Chan AP. **Predicting the Functional Effect of Amino Acid Substitutions and Indels.** PLoS One. 2012; 7(10).

Coleman, M. (1990). Is classical Rett syndrome ever present in males? Brain and Development, 12, 31–32.

Cornett, J., Cao, F., Wang, C.-E., Ross, C. A., Bates, G. P., Li, S.-H., Li, X.-J. **Polyglutamine expansion of huntingtin impairs its nuclear export.** Nature Genet. 2005.PubMed.

COSTA, Marcella Motta da. **Duplicação do gene MECP2 em meninos com deficiência intelectual.** 2016. 57 f., il. Dissertação (Mestrado em Ciências da Saúde)—Universidade de Brasília, Brasília, 2016.

Criteria for Rett Syndrome. **Ann Neurol.** v.23,1988.

Dayalu P, Albin RL. **Huntington disease: pathogenesis and treatment.** Neurol Clin. 2015 Feb; Elsevier.

Duijn E, Kingma EM, van der Mast RC 2007. **Psychopathology in verified Huntington's disease gene carriers.** J Neuropsychiatry Clin Neurosci.

Durr, A; Gargiulo, M; Feingold, J novembro de 2012. **The presymptomatic phase of Huntington disease.** Revue neurologique.

Futter, M., Diekmann, H., Schoenmakers, E., Sadiq, O., Chatterjee, K., Rubinsztein, D. C. **Wild-type but not mutant huntingtin modulates the transcriptional activity of liver X receptors.** J. Med. Genet. 2009.

Guo, Q. , Cheng, J. , Seefelder, M. , Engler, T., Pfeifer, G. , Oeckl, P. , Otto, M., Moser, F., Maurer, M. , Pautsch, A. , Baumeister, W. , Fernandez-Busnadiego, R. , Kochanek, S. **The cryo-electron microscopy structure of huntingtin** (2018) Nature, PubMed.

Hagberg, B., Hanefeld, F., Percy, A., Skjeldal, O. **An update on clinically applicable diagnostic criteria in Rett syndrome: comments to Rett syndrome clinical criteria consensus panel satellite to European Paediatric Neurology Society Meeting, Baden Baden, Germany, 11 September 2001.** Europ. J. Paediat. Neurol. 6: 293-297, 2002.

Hugo Verli. **Bioinformática da Biologia à flexibilidade molecular / 1º. ed. - São Paulo : SBBq, 2014. 282 p. : il.1. Bioinformática 2. Biologia Molecular.**

Kegel, K. B., Meloni, A. R., Yi, Y., Kim, Y. J., Doyle, E., Cuiffo, B. G., Sapp, E., Wang, Y., Qin, Z.-H., Chen, J. D., Nevins, J. R., Aronin, N., DiFiglia, M. **Huntingtin is present in the nucleus, interacts with the transcriptional corepressor C-terminal binding protein, and represses transcription.** J. Biol. Chem. 2002.PubMed.

KREBS, B. B.; DE MESQUITA, J. F. **Amyotrophic Lateral Sclerosis Type 20 - In silico Analysis and Molecular Dynamics Simulation of hnRNPA1.** PloS one, v. 11, n. 7, jul. 2016.

Leila Schuindt Monnerat, Aline dos Santos Moreira, Maria Carolina Viana Alves, Cibele Rodrigues Bonvicino, Fernando Regla Vargasad - Brain and Development - **Identification and characterization of novel sequence**

variations in MECP2 gene in Rett syndrome patients – Elsevier; Volume 32, Issue 10, November 2010, Pages 843-848 Brain and Development.

LOPES, F. et al. **Identification of novel genetic causes of Rett syndrome-like phenotypes.** Journal of medical genetics, v. 53, n. 3, p. 190–9, mar. 2016.

Matthias Michalek,[^] Evgeniy S. Salnikov,[^] and Burkhard Bechinger* **Structure and Topology of the Huntingtin 1–17 Membrane Anchor by a Combined Solution and Solid-State NMR Approach Biophys J.** 2013 Aug 6.

McFarland KN ,Huizenga MN ,Darnell SB ,Sangrey GR ,Berezovska O ,Cha J-HJ ,Outeiro TF ,Sadri-Vakili G. **MECP2: a new Huntingtin interactor.** Hum Mol Genet 2014 .

Neul, JL, Kaufmann, WE, Glaze, DG, Christodoulou, J., Clarke, AJ, Bahi-Buisson, N., Leonard, H., Bailey, MES, Schanen, NC, Zappella, M., Renieri, A. , Huppke, P., Percy, AK Rett síndrome : **revista critérios de diagnóstico e nomenclatura.** Ann. Neurol. 68: 944-950, 2010.

Ng PC, Henikoff S. **SIFT: Predicting amino acid changes that affect protein function.** Nucleic Acids Res.2003.

Philippart, M. (1990). The Rett syndrome in males.Brain and Development, 12, 33–36.

Rett A. Uber ein eigenartiges hirnatrophisches Syndrom bei Hyperammonämie in kindersalter. Wien Med Wochenschr 1966.

RODAN, L. H. et al. **A novel neurodevelopmental disorder associated with compoundheterozygous variants in the huntingtin gene.** European journal of human genetics : EJHG, v. 24, n. 12, p. 1826–1827, 2016.

Rosemberg GS, Arita FN, Campos C. **A Brazilian girl with the Rett syndrome.** Brain Dev 1986.

SCHWARTZMAN JS, ANGELA MARIA COSTA DE SOUZA, GUIDO FAIWICHOW, LUIZ HENRIQUE HERCOWITZ - **Rett syndrome phenotype in XXY karyotype: case report** - Arq. Neuro-Psiquiatr. v.56 n.4 São Paulo dez. 1998.

Schwartzman JS. **Rett syndrome** Programa de Pós-graduação em Distúrbios do Desenvolvimento da Universidade Presbiteriana Mackenzie. São Paulo, SP, Brasil.

Schwartzman JS. **Síndrome de Rett**. Temas sobre Desenvolvimento 1990.

THE RETT SYNDROME DIAGNOSTIC CRITERIA WORK GROUP. Diagnostic Topcu, M., Akyerli, C., Sayi, A., Toruner, GA, Kocoglu, SR, Cimbis, M., Ozcelik, T. **Mosaismo somático para uma mutação MECP2 associada à síndrome de Rett clássica em um menino** . Europ. J. Hum. Genet. 10: 77-81, 2002.

Tugba G Kucukkal,\$ Marharyta Petukh,\$ Lin Li, and Emil Alexov. **Structural and Physico-Chemical Effects of Disease and Non-Disease nsSNPs on Proteins**. Author manuscript; Curr Opin Struct Biol, 2016 Jun.

UNIPROT. Activities at the Universal Protein Resource (UniProt). Nucleic acids research, v. 42, n. 1, p. D191-8, jan. 2015.

Yvan Saeys, Iñaki Inza,Pedro Larrañaga **A review of resource selection techniques in bioinformatics** Bioinformatics , Volume 23, Edição 19, 1 de outubro de 2007, Páginas 2507–2517, Publicados: 24 de agosto de 2007.

Zahraa Haidar MSc Nadine Jalkh PhD Sandra Corbani MSc Joelle Abou-Ghoch MSc Ali Fawaz MD Cybel Mehawej PhD Eliane Chouery PhD **A Homozygous Splicing Mutation in PDE2A in a Family with Atypical Rett Syndrome**. Movement disorders 20 March 2020.

ZHANG, Y.; SKOLNICK, J. **TM-align: a protein structure alignment algorithm based on the TM-score.** Nucleic acids research, v. 33, n. 7, jan. 2005.

Zuccato, C., Ciammola, A., Rigamonti, D., Leavitt, B. R., Goffredo, D., Conti, L., MacDonald, M. E., Friedlander, R. M., Silani, V., Hayden, M. R., Timmusk, T., Sipione, S., Cattaneo, E. **Loss of huntingtin-mediated BDNF gene transcription in Huntington's disease.** Science, 2001.PubMed.