

研究タイトル	疾患原因となるアミノ酸変異の解析		
研究カテゴリー	計算生物学・バイオインフォマティクス		
学校名	三田国際学園高等学校		
都道府県	東京都		
研究者氏名	角野 陽奈美		
研究者(代表者)学年	2年(高校・高専)		

研究の要約

ヒトが疾患を発症する際、遺伝的要因が大きく作用するものを遺伝疾患と呼ぶ。その発症メカニズムの一つに、遺伝子の変異およびアミノ酸の置換(以後「アミノ酸変異」)によるタンパク質の構造、機能変化がある。しかし、アミノ酸変異と疾患発症の関係については不明な点が多い。本研究の目的は未知のアミノ酸変異から疾患発症を予測する機械学習モデルを構築することである。

まず、タンパク質を4つの局在先で分類し、それぞれの局在先においてアミノ酸変異がヒトに与える影響について検討した。続いて、局在先の情報などをトレーニングデータとし、XGBoost法を用いて未知のアミノ酸変異から疾患発症を予測する機械学習モデルを構築した。さらに高い精度の予測を行うため、DSSPから算出した変異場所の二次構造、露出表面積、保存性などいくつかのパラメータを追加し、再度予測モデルを構築した。

核では疾患との低い関連、ミトコンドリアでは疾患との高い関連がみられ、タンパク質を局在先で分類すると疾患との関連が異なる可能性が示唆された。1回目の予測モデルでは、約77%の未知のアミノ酸変異に対する疾患発症の有無が正しく予測された。2回目の予測モデルでは約82%の精度だった。DSSPを使用したことによるデータ数の減少や偏りが予測精度に影響した可能性があるため、今後は立体構造情報を用いない予測モデルをブラッシュアップしていく必要もあると考えた。

●確認事項

研究に用いているもの (人間、脊椎動物、微生物、組み換えDNA、細胞組織、どれも用いていない)	どれも用いていない
大学・研究機関などでの実験や装置使用があるか	いいえ(使用していない)
昨年までの研究からの継続研究か	はい(継続研究である)