

令和 3 年 6 月 23 日現在

機関番号：34204

研究種目：若手研究

研究期間：2018～2020

課題番号：18K18154

研究課題名（和文）タンパク質高次構造に基づくヒト意義不明バリエーションを評価する手法の開発

研究課題名（英文）Developing a method for evaluating variants of uncertain significance in humans based on protein structures

研究代表者

土方 敦司 (Hijikata, Atsushi)

長浜バイオ大学・バイオサイエンス学部・プロジェクト特任講師

研究者番号：80415273

交付決定額（研究期間全体）：（直接経費） 3,200,000円

研究成果の概要（和文）：本研究の目的は、タンパク質コード領域に見つかる意義不明なミスセンスバリエーションが、そのタンパク質の分子機能へ与える影響を評価するための手法を開発することである。タンパク質立体構造における各アミノ酸残基周辺の空間配置情報に着目し、疾患関連変異と中立変異とを比較したところ有意な差異を見出した。また、異なるアミノ酸残基の空間配置が類似している場合は、疾患関連変異よりも中立変異の場合に多いことがわかった。この結果より、空間配置を比較することで、その構造環境におけるアミノ酸置換の適応度を評価できる可能性が示唆された。

研究成果の学術的意義や社会的意義

がんの治療法選択や希少疾患の遺伝子診断においてクリニカルシーケンスが進展しているが、個人のゲノム上にはその意義が不明なバリエーションが多数存在しており、ゲノムバリエーションと疾患などの表現型との関連性をどのように評価するかが課題となっている。本研究によって開発した手法をさらに発展させることで、これらのバリエーションの分子機能への影響をより正確に評価することができると考えられ、正しく理解することにつながると考えられる。

研究成果の概要（英文）：In this research project, I tried to develop a method to evaluate the functional impacts of missense variants without certain clinical significance. For this purpose, I analyzed a spatial arrangement of residues surrounding a reference amino acid residue on a protein structure and found that different tendencies in those environments between the residues with known pathogenic and neutral variants. I also found a trend that the structural environments between amino acid replacement observed in neutral variants showed a certain similarity compared to those in pathogenic variants. These results suggest that the structural environment might be useful to evaluate the functional impacts of variants with uncertain significance.

研究分野：構造バイオインフォマティクス

キーワード：ゲノムバリエーション タンパク質立体構造 ミスセンス変異 アミノ酸残基 局所構造

1. 研究開始当初の背景

ヒトゲノム上の有害変異は、時として致死的であったり、疾患を引き起こしたりする。特定の遺伝子上的変異が原因となって引き起こされる、ヒト単一遺伝子疾患は、生物多様性の負の産物であり、これまでに7000種類以上の疾患が知られている。遺伝子疾患に関連する既知の疾患関連変異の約90%は、ゲノム全体の3%に満たないタンパク質コード領域に見つかっていることから、疾患等の表現型に関係する変異の多くは、タンパク質の機能と密接に結びついていると考えられる。そのため、ゲノム上の変異と疾患との関係を理解するためには、変異によるアミノ酸の変化がタンパク質の機能あるいは構造に与える影響の詳細を理解することが必要不可欠である。

近年の次世代シーケンサーの普及によって、ヒト個人のゲノム配列決定が容易となり、これまで原因が不明であった遺伝子性疾患の責任遺伝子とその疾患関連変異を特定できるようになってきた。しかし一方で、健常人においても数百種類の有害な変異が存在することや、非常に頻度の低いレアバリエーションについては、中立突然変異との区別が容易ではないなど、変異と疾患との関係を、塩基配列あるいはアミノ酸配列のみから説明することが困難である場合が多い。こうした変異は、意義不明バリエーション (Variants of Unknown Significance; VUS) として知られ、臨床現場において遺伝子検査を行う上での大きな障壁の一つとなっている。

遺伝子変異の影響を予測するツールは、これまでに数多く開発されている。それらの多くは、アミノ酸配列の進化的な保存性を考慮したもの (SIFT, Provean, FATHMM, MutationAssessor) か、保存性やその他の様々な特徴を機械学習法によって学習させたもの (PolyPhen2, SNAP, VEST, MutPred) である。しかし、ほとんどのツールは、疾患に関係するかどうかを判定するのみであり、変異がタンパク質の機能または構造に与える影響の詳細については答えてはくれない。また、疾患変異のアミノ酸残基については、代表者らを含む様々なグループによって、進化的保存性、タンパク質二次構造及びタンパク質立体構造上の溶媒露出度等との関連性が指摘されているが、変異残基周囲の空間的構造環境という視点で研究された例はまだない。

2. 研究の目的

本研究の目的は、アミノ酸残基のタンパク質立体構造上における空間的な構造環境に着目し、その残基における疾患関連変異が及ぼす分子機能への影響との関連性を明らかにすることで、意義の解釈が困難である VUS の分子機能及び構造への影響について評価するための手法を開発することである。疾患関連変異を持つアミノ酸残基とその周辺の空間的構造環境と、変異の分子機能及び構造への影響に関する実験データとを統合化することによって、遺伝子検査等の臨床現場における課題解決だけでなく、ゲノム変化による生物多様性メカニズムを理解するための基礎とする。

3. 研究の方法

(1) 疾患関連変異アミノ酸残基のタンパク質立体構造中における空間的環境特徴の抽出

疾患との関連が明らかとなっている疾患関連変異について、変異残基を含むタンパク質の立体構造データに基づき、そのアミノ酸残基を取り囲む空間的な構造環境(どの分子のどの残基がどこに位置するか)を定量的に記述するための方法の開発を行った。アミノ酸残基1つの構造情報を、炭素から炭素への原子座標がなすベクトルとして表現し、対象アミノ酸残基ベクトルと、その周囲のアミノ酸残基ベクトルの相対的な空間配置をデータセットとして記述する空間的環境特徴記述子 (LSED, Local Structural Environment Descriptor) を作成した。周囲に別のタンパク質分子やタンパク質以外の分子(例えば核酸や糖)があった場合についても同様にして基準原子座標系におけるベクトルに変換して記述した。

(2) 変異によるタンパク質機能・構造への影響に関する情報の文献等からの抽出

疾患関連変異によるタンパク質の機能あるいは構造への影響として、(a) フォールディングやタンパク質の熱安定性、(b) タンパク質の細胞内局在性、(c) 分子間相互作用による高次構造形成性、(d) 酵素反応やDNA結合等の分子機能性、(e) リン酸化等の翻訳後修飾による分子機能制御性、(f) タンパク質の分解などが挙げられる。既存の研究によって、特定のヒト遺伝子に関して、実験によって変異とその分子の機能・構造への影響を網羅的に解析されていて論文として発表されているデータを論文 Supplementary Table 等から抽出し、統一されたフォーマットでデータを整理・統合したデータベースを作成した。

(3) データの統合化及び病因変異における空間的構造環境特徴と機能・構造的影響との関連性の解析

上記(1)及び(2)によって作成したデータを統合化して、相互に参照可能なデータベースを作

成した。このデータベースに基づき、疾患変異の立体構造環境の特徴と、変異のタンパク質分子機能及び構造への影響との間に関連性があるかについての検討を行った。

(4) 残基の空間的環境特徴を用いた VUS の機能・構造的影響の推定法の検討

任意のバリエーションについて、上記(1)と同様に、LSED を作成した。それらを既知変異のベクトルデータセットと比較し、類似度スコアを算出する機能を上記(3)で作成したデータベース上に実装した。このシステムを用いたブラインドテストにより、既知病因変異の影響を正しく判定できるかの評価を行った。

4. 研究成果

(1) 立体構造中の任意のアミノ酸残基の空間的環境特徴記述子(LSED)の相互比較の手法の開発
タンパク質立体構造から、特定のアミノ酸残基(基準アミノ酸残基)の周囲のアミノ酸残基の空間配置を、炭素と炭素がなすベクトル情報として表現して LSED を抽出するプログラムを開発した。並行して、アミノ酸残基の LSED を相互の類似度を算出するアルゴリズムを実装したプログラムを作成した。このことにより、任意の2残基の LSED の類似度を定量的に比較することが可能となった。

(2) LSED の網羅的解析

次に、ヒトの遺伝子疾患の原因タンパク質 4,618 個のうち、立体構造が既知である 2,655 個のタンパク質に着目した解析を行った。これらのタンパク質に含まれる既知の疾患ミスセンス変異アミノ酸残基(10,217 残基)について、上記のプログラムを用いて LSED を抽出した。基準アミノ酸残基が同じアミノ酸であるデータについて総当たりで類似度を算出した。その結果、全体の約 15%のアミノ酸残基(1,626 個)の LSED について、進化的な類縁関係がないタンパク質においても類似性があることがわかった。このことが意味することは、類縁フォールドの垣根を超えて共通した分子局所構造因子の存在を示唆するものである。

(3) LSED と疾患関連変異及び中立変異との関連性の解析

次に、上記で開発した手法を用いて、疾患変異残基と、疾患とは無関係な中立変異残基について、LSED を相互に比較した。多くの場合において、異なる特徴を持っていることがわかった。また同様の方法で、VUS 残基(18,137 残基)の空間的環境特徴記述子とも比較したところ、疾患関連変異残基と非同義置換残基の中間の性質を持っている場合が多いことがわかった。

また、疾患変異残基の LSED 及び中立変異残基の LSED を総当たりで比較したところ、疾患疾患ペアまたは中立 中立ペアの類似性の方が、疾患 中立ペアと比べて統計的に有意に高いことがわかった。

(4) LSED の類似性に基づく VUS の影響評価手法の検討

次に、網羅的な変異機能解析の実験データと組み合わせた解析を行った。ここでは、以下のような仮定を置いて解析した。すなわち、もし対象アミノ酸残基の LSED が、置換先のアミノ酸を基準とする LSED と類似していれば、そのアミノ酸置換は許容される、と考える。がん抑制遺伝子 TP53 及び PTEN の構造既知の各アミノ酸残基(それぞれ 177、313 アミノ酸残基)について、LSED を抽出し、上記のデータセットを用いて網羅的な類似性検索を行なった。その結果、想定通り、置換先のアミノ酸残基 LSED と高い類似性スコアを示した場合、そのアミノ酸置換は機能に影響を与えない場合が有意に多い結果となった。

本研究の結果は、VUS を評価する上で、LSED を用いた解析が有効であるだけでなく、タンパク質の構造 機能相関を理解する上でも重要な知見であると考えられる。しかしながら、本研究期間内で行った解析では、2種類の遺伝子のみを対象とした結果であり、十分な検討がなされたとは言い難い。また LSED と機能との関連性についてもまだ検討の余地が残されている。今後さらに検討を重ねていくことで、予測手法をさらに洗練されたものとし、VUS の生物学的意義についての理解を深める手法へとつなげていきたいと考えている。

5. 主な発表論文等

〔雑誌論文〕 計3件（うち査読付論文 2件/うち国際共著 0件/うちオープンアクセス 3件）

1. 著者名 Hijikata Atsushi	4. 巻 2018
2. 論文標題 Developing a new method to predict the effect of the missense mutations on the protein function and phenotypes	5. 発行年 2018年
3. 雑誌名 Impact	6. 最初と最後の頁 35～37
掲載論文のDOI（デジタルオブジェクト識別子） 10.21820/23987073.2018.12.35	査読の有無 無
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 -

1. 著者名 Takasawa Kei, Tsuji-Hosokawa Atsumi, Takishima Shigeru, Wada Yasunori, Nagasaki Keisuke, Dateki Sumito, Numakura Chikahiko, Hijikata Atsushi, Shirai Tsuyoshi, Kashimada Kenichi, Morio Tomohiro	4. 巻 11
2. 論文標題 Clinical characteristics of adolescent cases with Type A insulin resistance syndrome caused by heterozygous mutations in the -subunit of the insulin receptor (INSR) gene	5. 発行年 2019年
3. 雑誌名 Journal of Diabetes	6. 最初と最後の頁 46～54
掲載論文のDOI（デジタルオブジェクト識別子） 10.1111/1753-0407.12797	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 -

1. 著者名 Sutani Akito, Shima Hirohito, Hijikata Atsushi, Hosokawa Susumu, Katoh-Fukui Yuko, Takasawa Kei, Suzuki Erina, Doi Shozaburo, Shirai Tsuyoshi, Morio Tomohiro, Fukami Maki, Kashimada Kenichi	4. 巻 63
2. 論文標題 WDR11 is another causative gene for coloboma, cardiac anomaly and growth retardation in 10q26 deletion syndrome	5. 発行年 2020年
3. 雑誌名 European Journal of Medical Genetics	6. 最初と最後の頁 103626～103626
掲載論文のDOI（デジタルオブジェクト識別子） 10.1016/j.ejmg.2019.01.016	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 -

〔学会発表〕 計5件（うち招待講演 1件/うち国際学会 0件）

1. 発表者名 土方 敦司、塩生 真史、白井 剛
2. 発表標題 タンパク質高次構造情報に基づくヒトミスセンス変異の分子機能への影響を予測する手法の開発
3. 学会等名 第19回日本蛋白質科学会年会 第71回日本細胞生物学会大会 合同年次大会
4. 発表年 2019年

1. 発表者名 Atsushi Hijikata, Masafumi Shionyu, Tsuyoshi Shirai
2. 発表標題 A local structural environment descriptor towards evaluating impact of rare variants in humans on protein structures and functions
3. 学会等名 第57回日本生物物理学会年会
4. 発表年 2019年

1. 発表者名 土方 敦司、塩生 真史、白井 剛
2. 発表標題 タンパク質高次構造に基づくヒトミスセンスバリエーションの分子機能への影響を予測する手法の開発
3. 学会等名 第42回日本分子生物学会年会
4. 発表年 2019年

1. 発表者名 土方敦司、塩生真史、白井剛
2. 発表標題 Development of a method for predicting pathogenicity of missense variants incorporating supramolecular structural information
3. 学会等名 第56回日本生物物理学会年会
4. 発表年 2018年

1. 発表者名 土方敦司
2. 発表標題 たんぱく質の高次構造から読み解くゲノム変異と疾患表現型との関係
3. 学会等名 第17回ゲノム・オミックス連携推進セミナー（招待講演）
4. 発表年 2019年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
--	---------------------------	-----------------------	----

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関
---------	---------