

厚生労働科学研究費補助金
難治性疾患等克服研究事業（難治性疾患等克服研究事業）
遺伝学的検査の実施拠点の在り方に関する研究
分担研究報告書

患者登録（レジストリー）変異データベース構築に関する課題

研究分担者 松田 文彦 京都大学大学院医学研究科・教授
研究分担者 松原 洋一 国立成育医療研究センター・所長

研究要旨

遺伝子診断体制の確立にあたり、同定された変異に適切な解釈を付与するためには、患者の臨床情報や遺伝子変異の生物学的意義等が集約されたデータベースを参照する必要がある。厚生労働省科学研究費「難病・がん等の疾患分野の医療の実用化研究事業」において遺伝子解析拠点研究班（5拠点）が協力して構築した「日本人の遺伝子変異データベース・疾患変異データベース」は、我が国における遺伝子診断事業に不可欠の日本人のゲノム変異情報を蓄積している。多様な疾患の遺伝子診断結果を本データベースへ継続的に蓄積し、遺伝子変異情報を集約・一元化することにより、高精度かつ標準化された手法での遺伝子診断が実施され、疾患の再分類が進み、適切な医療体制が確立されると期待される。このような体制の構築や維持・運営のためには、解析拠点施設が中心となりゲノムインフォマティクス人材の育成に尽力する必要がある。

A. 研究目的

遺伝学的検査体制の構築にあたり、遺伝子変異に応じた精度の高い診断、迅速な治療方針の決定による質の高い「個の医療」を実現するためには、遺伝子変異情報を集約し、標準化・一元管理のうえ参照可能にする情報基盤の整備が必要不可欠である。本課題では、データベースおよびゲノムインフォマティクスの現状について調査し、遺伝子診断体制に必須の情報基盤構築に関する課題について考察する。

B. 研究方法

厚生労働科学研究費「難治性疾患克服研究事業」の遺伝子解析5拠点（京都大学、東京大学、東北大学、横浜市立大学、国立成育医療研究センター）が連携し「日本人の遺伝子変異データベース・疾患変異データベース」を構築した。そして2013年11月より研究者コミュニティに公開している（<http://www.genome.med.kyoto-u.ac.jp/SnpDB>）。本データベースは、日本人を対象としたゲノム解析研究において有用なゲノム変異情報を網羅しており、遺伝子診断体制に必要な情報基盤としては、最適なデータベースである。本研究課題では、当該データベースの現状を報告し、遺伝子診断体制の実地に際した情報集約の意義と課題

について報告する。

（倫理面への配慮）

データベースの構築に際し、研究代表者、分担者はそれぞれの機関において審査を受け、適切に実施されている。連携機関とのデータのやりとりは、匿名化された情報のみ扱われている。

C. 研究結果

日本人の遺伝子変異データベース

遺伝子診断においては、同定された遺伝子変異に対する解釈が最も重要となる。この解釈においては、大規模な健常者集団のゲノム変異情報との比較や、疾患毎の変異を集約したデータベースの参照が不可欠である。ゲノム多様性、疾患関連変異は、集団に固有のものが少なくないことがわかってきており、我が国での遺伝子診断体制の確立には日本人集団のデータベース構築が不可欠である。先に公開された1,208名の健常者日本人集団の「日本人の遺伝子変異データベース」は、5つの次世代シーケンサー解析拠点が緊密に連携して構築した唯一の日本人集団のデータベースであり、研究者コミュニティがさまざまな研究、診療に活用することができるわが国の生命科学研究における基盤的役割を果たしている。

本データベースは、公開から 3 ヶ月で 717,654 回のアクセスを得ており、疾患ゲノム解析への関心の高さ、含まれるゲノム情報の有用性が証明された。今後、様々な疾患でエクソーム関連解析が増加すると予測されるが、そのような研究の基盤として、さらに検体数・情報量をふやし、データベースの価値を高めることが重要である。

以下、公開情報の概要を列挙する。

- ・ 1,208 名のエクソーム解析結果より算出した 288,025 座位の日本人のエクソン領域における遺伝子多型頻度情報
- ・ 3,248 名の一塩基多型タイピング解析結果より算出した 1,794,196 座位の日本人のゲノム多型頻度情報
- ・ 300 名の発現アレイ解析とエクソーム解析情報を組み合わせた 21,755 遺伝子における網羅的発現関連解析 (eQTL) 情報

疾患変異データベース

疾患と関連する遺伝子変異のデータベースは、個別の疾患の小規模データベースは存在するが、幅広く多くの疾患を網羅したデータベースは存在しない。次世代シーケンサーを用いた遺伝子診断を拠点で一元化して実施することは、データが集積されること、統一プロトコルで質の担保されたデータの産生を可能とすることなど、このような変異データベースの構築に大きく寄与すると考えられる。また、疾患関連変異データベースが、我が国での標準として研究者・医療者に長期にわたって利活用されるには、拠点の解析結果に加えて個別研究で見出される疾患関連変異も、本データベースへの登録を促すことが重要である。

国際的には、Human Variome Project という大きな動きがあり、全世界でヒトゲノム多様性情報のデータベースの構築が推進されている。このような国際的な試みに貢献していくためにも、次世代シーケンサー解析拠点が中心となり、Japan Node を設置し役割を担うことが求められる。

ゲノムインフォマティクス

次世代シーケンサーを用いた大規模ゲノム解析においては、ゲノムインフォマティクスの持つ役割が飛躍的に増大してきており、今後は、大規模臨床情報を含めた生命ビッグデータの研究を推進する学際的研究分野として発展することが予想される。次世代シーケンサーから産生される膨大

な情報からの配列情報の抽出、ヒトゲノム参照配列への整列、変異のコールなどの一連の解析、品質管理、さらに、病的意義を有する変異の探索、評価においては、さまざまなデータベースを参照すること、機能解析の予測など、重層的な解析が求められる。これらの解析を実施するためには、大規模並列処理ができる計算サーバーを備え、インフォマティクスの研究者が担当する研究体制が必要であり、このような役割を次世代シーケンサー解析拠点が担う必要がある。また、医学系の情報と、大規模ゲノム情報を同時に扱い、OJT(On the Job Training)が可能な環境が、人材不足が指摘されているメディカルゲノムインフォマティクス分野の人材育成にとって極めて重要となり、このような環境の構築・提供が次世代シーケンサー解析拠点の役割となる。京都大学では、海外より著名な講師陣を招き、次世代シーケンサー解析のトレーニングコースを年一回実施しており、教育だけでなく実践技術を習得する環境を提供し、日本のバイオインフォマティクス全体のレベル向上に努めている。さらに、新規アルゴリズムの開発など、研究面で高いレベルを目指すことも大切であり、そのような研究の場を提供し、ゲノムインフォマティクスの深い知識と豊富な経験を持つ研究者を育成する機能が、次世代シーケンサー解析拠点の重要な役割となる。

京都大学トレーニングコース概要

- ・「第 1 回次世代シーケンス解析トレーニングコース」2013 年 1 月 15 日～18 日、京都大学大学院医学研究科 芝蘭会館
- ・「第 2 回次世代シーケンス解析トレーニングコース」2014 年 3 月 10 日～12 日、京都大学大学院医学研究科 南部総合研究 1 号館

D. 考察

データエントリー

日本人の遺伝子変異データベースに登録されている 288,025 個の変異のうち、156,622 個 (54.4%) は、世界の公開データベースには存在しない日本人特異的な新たな変異であった。これらの変異の約 60% はアミノ酸変化を伴い、タンパク質の機能に影響を与える変異であったことより、これらの日本人特異的ゲノム変異は、遺伝子の機能に関わる重要な変異の可能性が高く、疾患に関連する遺伝子変異の効率的な探索

や診断に有用と考えられる。

これらの情報に加えて、遺伝子変異と遺伝子発現量の関連解析 (eQTL) 情報も記載されており、新規に同定された疾患関連変異の機能的役割を解釈するうえで重要な情報基盤である。

情報基盤整備により期待される成果

- 1) 多種の難病の遺伝子診断体制が構築され、遺伝子診断における標準化が進む。
- 2) 遺伝子産物の構造や機能に影響を与える遺伝的変異が収集され、難病の発症機構の解明やゲノム変異に基づいた難病の再分類に役立つ。
- 3) 本データベースを参照することにより、遺伝子診断を行なった患者で同定された遺伝子変異が疾患に関連するかどうかの検証が容易に行える。
- 4) 疾患遺伝子情報のみならず、疾患背景や治療経過などの情報の収集により疾患の全体像の理解が進み、適切な医療体制を敷く基礎となるデータを提供できる。

今後の課題

・情報の集約・一元化

疾患変異データベースとして機能するための最優先課題は、既知の全ての変異情報の集約・一元化である。これにより、解析結果への適切な解釈が可能となる。

・登録の推進

遺伝子診断体制の実用化と登録の義務化により、変異情報が集約されていくシステム構築が望まれる。各研究機関に死蔵されている過去のデータ等を抽出するためには、登録への働きかけも必要とされる。また、変異の解釈に有用な充実した臨床情報の提供や科学雑誌との連携により登録のインセンティブを確保することも重要である。

・登録支援

各施設で同定された変異は、必ずしも単一のプラットフォームで解析されたものとは限らず、変異の表記法も統一されていない。そのような場合にも対応できるような登録支援インターフェイスを実装する。

・情報の充実化

諸外国のデータベースとの連携や、希少疾患関連変異の探索に対応するため検体数や情報量の拡大をおこなう。

・品質管理

登録に際し、品質管理の標準化・プロト

コルの策定、また、信頼度の指標・確認実験手法等の情報も付加する。

・患者登録

生体試料バンク事業と連携をとる。実際の運用には、コホート事業で構築し実績のある臨床情報管理システムを利用する。

・長期的運用・維持

利用価値の高いデータベースを構築しても、資金提供のないままの運用・維持は困難であるため、事業化による長期にわたる資金的援助が強く望まれる。また、解析を専門とする人材の育成が必要である。

E . 結論

迅速かつ正確な遺伝子診断体制を確立するためには、変異情報を集約・一元化し、様々な医療機関から登録・参照可能にする情報基盤の整備が緊要の課題である。現在、遺伝子解析 5 拠点 が連携して構築し、維持・運営している日本人の遺伝子変異・疾患関連変異データベースは、遺伝子診断に最適な情報基盤を構築しつつあり、観測された変異に対する解釈付与において極めて重要な役割を果たすことが期待される。今後データベースの内容をさらに充実したものとし、長期にわたって継続的な運用を行うためには、変異登録のインセンティブを確保し、ゲノムインフォマティクス人材を育成する環境を提供することが拠点機関の重要な役割である。

F . 健康危険情報

なし

G . 研究発表

【論文発表】

1. Okada, Y., Wu, D., Trynka, G., Raj, T., Terao, C., Ikari, K., Kochi, Y., Ohmura, K., Suzuki, A., Yoshida, S., Graham, R. R., Manoharan, A., Ortmann, W., Bhangale, T., Denny, J. C., Carroll, R. J., Eyler, A. E., Greenberg, J. D., Kremer, J. M., Pappas, D. A., Jiang, L., Yin, J., Ye, L., Su, D. F., Yang, J., Xie, G., Keystone, E., Westra, H. J., Esko, T., Metspalu, A., Zhou, X., Gupta, N., Mirel, D., Stahl, E. A., Diogo, D., Cui, J., Liao, K., Guo, M. H., Myouzen, K., Kawaguchi, T., Coenen, M. J. H., van Riel, P. L. C. M., van de Laar, M. A. F. J., Guchelaar, H. J., Huizinga, T. W. J.,

- Dieude, P., Mariette, X., Bridges Jr, S. L., Zhernakova, A., Toes, R. E. M., Tak, P. P., Miceli-Richard, C., Bang, S. Y., Lee, H. S., Martin, J., Gonzalez-Gay, M. A., Rodriguez-Rodriguez, L., Rantapaa-Dahlqvist, S., Arlestig, L., Choi, H. K., Kamatani, Y., Galan, P., Lathrop, M., the RACI consortium, the GARNET consortium, Eyre, S., Bowes, J., Barton, A., de Vries, N., Moreland, L. W., Criswell, L. A., Karlson, E. W., Taniguchi, A., Yamada, R., Kubo, M., Liu, J. S., Bae, S. C., Worthington, J., Padyukov, L., Klareskog, L., Gregersen, P. K., Raychaudhuri, S., Stranger, B. E., De Jager, P. L., Franke, L., Visscher, P. M., Brown, M. A., Yamanaka, H., Mimori, T., Takahashi, A., Xu, H., Behrens, T. W., Siminovitch, K. A., Momohara, S., Matsuda, F., Yamamoto, K. and Plenge, R. M. (2014) Genetics of rheumatoid arthritis contributes to biology and drug discovery. *Nature* **506**, 376-381.
2. Tanaka, K., Terao, C., Ohmura, K., Takahashi, M., Nakashima, R., Imura, Y., Yoshifuji, H., Yukawa, N., Usui, T., Fujii, T., Mimori, T. and Matsuda, F. (2014) Significant association between *CYP3A5* polymorphism and blood concentration of tacrolimus in patients with connective tissue diseases. *J. Hum. Genet.* **59**, 107-109.
 3. Yamakawa, N., Fujimoto, M., Kawabata, D., Terao, C., Nishikori, M., Nakashima, R., Imura, Y., Yukawa, N., Yoshifuji, H., Ohmura, K., Fujii, T., Kitano, T., Kondo, T., Yurugi, K., Miura, Y., Maekawa, T., Saji, S., Takaori-Kondo, A., Matsuda, F., Haga, H. and Mimori, T. (2014) A clinical, pathological and genetic characterization of methotrexate-associated lymphoproliferative disorders. *J. Rheumatol.* **41**, 293-299.
 4. Terao, C., Bayoumi, N., McKenzie, C. A., Zelenika, D., Muro, S., Mishima, M.; The Nagahama Cohort Research Group, Connell, J. M., Vickers, M. A., Lathrop, G. M., Farrall, M., Matsuda, F. and Keavney, B. D. (2013) Quantitative variation in plasma angiotensin-I converting enzyme activity shows allelic heterogeneity in the ABO blood group locus. *Ann. Hum. Genet.* **77**, 465-471.
 5. Terao, C., Yoshifuji, H., Ohmura, K., Murakami, K., Kawabata, D., Yurugi, K., Tazaki, J., Kinoshita, H., Kimura, A., Akizuki, M., Kawaguchi, Y., Yamanaka, H., Miura, Y., Maekawa, T., Saji, H., Mimori, T. and Matsuda, F. (2013) Association of Takayasu arteritis with HLA-B*67:01 and two amino acids in HLA-B protein. *Rheumatol. (Oxford)*. **52**, 1769-1774.
 6. Terao, C., Yoshifuji, H., Kimura, A., Matsumura, T., Ohmura, K., Takahashi, M., Shimizu, M., Kawaguchi, T., Chen, Z., Naruse, T. K., Sato-Otsubo, A., Ebana, Y., Maejima, Y., Kinoshita, H., Murakami, K., Kawabata, D., Wada, Y., Narita, I., Tazaki, J., Kawaguchi, Y., Yamanaka, H., Yurugi, K., Miura, Y., Maekawa, T., Ogawa, S., Komuro, K., Nagai, R., Yamada, R., Tabara, Y., Isobe, M., Mimori, T. and Matsuda, F. (2013) Two susceptibility loci to Takayasu arteritis reveal a synergistic role of the IL12B and HLA-B regions in a Japanese population. *Am. J. Hum. Genet.* **93**, 289-297.
 7. Plenge, R. M., Greenberg, J. D., Mangravite, L. M., Derry, J. M., Stahl, E. A., Coenen, M. J., Barton, A., Padyukov, L., Klareskog, L., Gregersen, P. K., Mariette, X., Moreland, L. W., Bridges, S. L. Jr, de Vries, N., Huizinga, T. W. Guchelaar, H. J., International Rheumatoid Arthritis Consortium (INTERACT), Friend, S. H. and Stolovitzky, G. (2013) Crowdsourcing genetic prediction of clinical utility in the Rheumatoid Arthritis Responder Challenge. *Nat. Genet.* **45**, 468-469.
 8. Sekiguchi, K., Maeda, T., Suenobu, S., Kunisaki, N., Shimizu, M., Kiyota, K., Handa, Y.S., Akiyoshi, K., Korematsu, S., Aoki, Y., Matsubara, Y. and Izumi, T. (2013) A transient myelodysplastic/myeloproliferative neoplasm in a patient with cardio-facio-cutaneous syndrome and a germline BRAF mutation. *Am. J. Med. Genet. A*, **161**, 2600-2603.
 9. Miyake, N., Koshimizu, E., Okamoto, N., Mizuno, S., Ogata, T., Nagai, T., Kosho,

T., Ohashi, H., Kato, M., Sasaki, G., Mabe, H., Watanabe, Y., Yoshino, M., Matsuishi, T., Takanashi, J., Shotelersuk, V., Tekin, M., Ochi, N., Kubota, M., Ito, N., Ihara, K., Hara, T., Tonoki, H., Ohta, T., Saito, K., Matsuo, M., Urano, M., Enokizono, T., Sato, A., Tanaka, H., Ogawa, A., Fujita, T., Hiraki, Y., Kitanaka, S., Matsubara, Y., Makita, T., Taguri, M., Nakashima, M., Tsurusaki, Y., Saito, H., Yoshiura, K., Matsumoto, N. and Niikawa, N. (2013) MLL2 and KDM6A mutations in patients with Kabuki syndrome. *Am. J. Med. Genet. A* **161**, 2234-2243.

【学会発表】

なし

H . 知的財産権の出願・登録状況 (予定を含む。)

1. 特許取得

なし

2. 実用新案登録

なし

3. その他

なし