

- Human Genetics, 2012
6. Kosaki R, Okuno N, Torii C, Kosaki K: Hartsfield syndrome and Xq 24 microduplication European Society of Human Genetics, 2011
 7. Kosaki K, Tanaka R, Kosaki R, Uchida C, Torii C, Ishi T, Sato T, Yoshihashi H. Wide phenotypic variability of Kabuki syndrome with MLL2 mutations. American Society of Human Genetics, 2011
 8. 鳴海洋子 仁科幸子 時光元温 青木洋子 小崎里華 涌井敬子 村田敏規 高田史男 古庄知己 福嶋義光 :先天性白内障家系におけるMAF遺伝子変異の同定 第58回 日本人類遺伝学会 2013.11.22
 9. 佐々木 愛子 藤田 秀樹 和田 誠司 小崎里華 堀川 玲子 左合 治彦:高齢妊娠を契機に羊水検査を行いX/XY 混合性性腺異形成と診断された3 例の臨床経過 第58回 日本人類遺伝学会 2013.11.22
 10. 藤田 秀樹 小崎里華:FGFR2 変異を確認したBeare-Stevenson 症候群の一例 第58回 日本人類遺伝学会 2013.11.22
 11. 小崎健次郎 小崎里華:先天異常症候群診断パネルの臨床応用 第3回 NGS現場の会 2013.9.4
 12. 藤田秀樹 小崎里華:20q端部欠失を認めたVATER連合に類似す多発奇形症例第53回 日本先天異常学会 2013.7.22
 13. 佐々木愛子 藤田秀樹 和田誠司 住江正大 杉林里佳 上出泰山 岡田朋美 太崎友紀子 藤村千鶴子 李紅蓮 小崎里華 左合治彦 :日本における出生前診断の現況 第37回 日本遺伝カウンセリング学会 2013 6.21
 14. 竹澤祐介 余谷暢之 石黒精 紙谷万里子 益田博司 師田信人 小崎里華 宮寄治 西村玄 阪井裕一:当センターで経験した点状軟骨異形成症の環軸亜脱臼に関する臨床的経過の検討 第116回 日本小児学会学術集会 2013.4.20
 15. 高木優樹 藤田秀樹 小崎里華 内木康博 鳴海覚志 三宅紀子 鶴崎美德 才津浩智 中島光子 松本直通 西村玄 長谷川奉延 :SERPINH1遺伝子変異による骨形成不全症の本邦初発例 第36回 日本小児遺伝学会学術集会 2013.4.17
 16. 藤田 秀樹 小崎里華 :染色体G分染法で48,XYYYと 49,XYYYYのモザイクを認めた一例 第36回 日本小児遺伝学会学術集会 2013.4.17.
 17. 岡田 朋美 大柴 葉子 佐々木 愛子 谷口 公介 杉林 里佳 住江 正大 和田 誠司 柿島 裕樹 小須賀 基通 小崎里華 小澤 伸晃 左合治彦:羊水検査においてFISH 法 (AneuVysion) とG-band 法で異なった核型結果が得られた2 症例 第57回日本人類遺伝学会 2012.10.26
 18. 藤田 秀樹 小崎里華: 当院で経験した未診断の発達遅滞・先天性多発奇形症例のアレイCGH 解析(続報) 第57回日本人類遺伝学会 2012.10.26
 19. 和田 友香 花井 彩江 佐々木 愛子 藤田 秀樹 小崎里華: 近四倍体と18 トリソミーのモザイクの1 女児例 第57回日本人類遺伝学会 2012.10.26
 20. 鳥居千春 丸岡亮 清水厚志 小崎里華 小崎健次郎:次世代シーケンサーを用いた先天奇形症候群の網羅的診断 第52回 日本先天異常学会 2012. 7.6
 21. 藤田秀樹 小崎里華 :アレイCGHで1q24-q25 領域に欠失を認めた精神運動発達遅滞児についての検討 第52回 日本先天異常学会 2012.7.6
 22. 三須久美子 桐林和代 佐谷秀行 鳥居千春 小崎里華 小崎健次郎:神経線維腫症1型の遺伝様式に関する患者家族の誤解の類型化 第36回 日本遺伝カウンセリング学会 2012 6.10
 23. 岡田朋美 佐々木愛子 黒田くみ子 上田英梨子 江川真希子 杉林里桂 住江正大 李紅蓮 藤田秀樹 小崎里華 左合治彦: 成育医療研究センターにおける周産期遺伝カウンセリング体制 第36回 日本遺伝カウンセリング学会 2012 6.9
 24. 武内俊樹 下郷幸子 山崎麻美 小崎里華 小崎健次郎 高橋高雄: L1CAM変異により発症した先天性水頭症とHirshsprung 病の合併例 第54回 日本小児神経学会総会 2012.4.19
 25. 柳橋達彦 小崎健次郎 小崎里華 吉橋博史 井原正博 高橋孝雄:Williams症候群における欠失範囲の大きさと発達遅滞の重症度との関係 第115回 日本小児学会学術集会 2012.4.22
 26. 一宮優子 石黒精 中館尚也 前川貴伸 小崎里華 藤田秀樹 阪井裕一 :TPO受容体作動薬が有効であった3p12-13欠失の難治性慢性ITP 第115回 日本小児学会学術集会 2012.4.22
 27. 藤田 秀樹 小崎里華 :アレイCGHで2p部分欠失を認めた自閉症患者についての検討 第35回 日本小児遺伝学会学術集会 2012.4. 19
 28. 李紅蓮 林聡 左合治彦 小崎里華:妊娠中に Trisomy18と診断された妊婦の意思決定 第35回 遺伝カウンセリング学会 2011
 29. 藤田秀樹 小崎里華 :閉症患者検体を用いたアレイCGH 解析による原因遺伝子の検討 第51回日本先天異常学会学術集会 2011.7.22
 30. 柳橋達彦 小崎健次郎 岡本伸彦 水野誠司 黒澤健司 小崎里華 :Rubinstein_Taybi 症候群の精神症状の経時的変化と治療可能性 第51回日本先天異常学会学術集会 2011.7.23
 31. 小崎里華 小崎健次郎 黒澤健司 岡本伸彦 水野誠司 高山真一郎:EEC 症候群 (Ectrodactyly, Ectodermal dysplasia Cleft

- lip_palate syndrome)の本邦実態調査 第51回日本先天異常学会学術集会 2011.7.23
32. 藤田秀樹 小崎里華 :G分染法でde novo均衡型と診断されたてんかん・発達遅滞症例のアレイCGH解析 日本小児遺伝学会学術集会 2011.8. 11
33. 柳橋達彦 小崎健次郎 岡本伸彦 水野誠司 黒澤健司 小崎里華. Rubinstein_Taybi 症候群のbehavioral pattern の経時的变化 日本小児遺伝学会学術集会 2011. 8.11
34. 田中竜馬 小崎健次郎 吉橋博史 小崎里華 高橋孝雄 :MLL2変異の確認されたカブキ症候群10例の臨床像 日本小児学会学術集会 2011.8. 13
35. 藤田秀樹 小須賀基道 奥山虎之 小崎里華 :当センターで経験したX連鎖重症複合免疫不全症(X-SCID)の出生前診断の4例 日本小児学会学術集会 2011.8. 13
36. 河合利尚 村山静子 新井勝大 小崎里華 奥山虎之 小野寺雅史 :慢性肉芽腫症における非感染性炎症疾患の検討 日本小児学会学術集会 2011.8. 13
37. 塚口裕康 仲里仁史 森貞直哉 服部元史 伊藤秀一 小崎里華 飯島誠球(体異形成)の全エクソーム解析 第56回日本人類遺伝学会 2011
38. 武内俊樹 柳橋達彦 土屋裕行 鳥居千春 林久美子 小崎里華 高橋孝雄 小崎健次郎:モザイク環状18番染色体によるPitt-Hopkins症候群の1例 第56回日本人類遺伝学会 2011
39. 仁科幸子 小崎里華 東範行 岡本伸彦 初川嘉一 黒澤健司 山根敬浩 水野誠司 都築欣一 小崎健次郎:CHD7遺伝子変異によるCHARGE症候群の眼合併症 第56回日本人類遺伝学会 2011
40. 佐々木愛子 鈴木朋 今野秀洋 住江正大 林聡 左合治彦 小崎里華 鏡雅代:upd14の出生前診断を行った1例 第56回日本人類遺伝学会 2011
41. 小崎里華 水野誠司 岡本伸彦 黒澤健司 小崎健次郎 高山 真一郎 EEC症候群における有病率の調査と実態調査研究 第56回日本人類遺伝学会 2011
42. 藤田秀樹 鳥居千春 小崎健次郎 小崎里華:当院で経験したAEC(Ankyloblepharon-Ectodermal dysplasia-Cleft lip and/or palate) syndromeの一例 第56回日本人類遺伝学会 2011
43. 小崎健次郎 星野健 小崎里華 高山真一郎 岡本伸彦 水野 誠司 黒澤健司:VATER連合の症状に関する全国調査とデータマイニング 第56回日本人類遺伝学会 2011
44. 岡田朋美 佐々木愛子 林聡 李紅蓮 藤田秀樹 小崎里華 左合 治彦 当センターにおける絨毛生検の検討 第56回日本人類遺伝学会 2011

45. 「先天奇形の疫学、分類と診断、染色体異常、出生前診断」 妊娠と薬情報センター 2011.2.11
46. 「Rubinstein-Taybi症候群 について」 ルビンステイン・テエイビ症候群 家族会 「こすもす」交流会 2011. 7. 23

G. 知的財産権の出願・登録状況(予定を含む。)

1. 特許取得
2. 実用新案登録
3. その他
なし

厚生労働科学研究費補助金
難病・がん等の疾患分野の医療の実用化研究事業（難病関係研究分野）
総合研究報告書

次世代シーケンサーによる解析法の最適化と標準化に関する研究

研究分担者 工藤 純 慶應義塾大学医学部遺伝子医学研究室 教授

研究要旨

次世代シーケンサーを用いた全ゲノム解読あるいは全翻訳領域解読（いわゆるエキソーム解析）は、未知の疾患原因遺伝子を探索する研究手段として確立されつつあり、われわれの研究班でも遺伝性疾患の新規原因遺伝子の探索や未診断例の原因究明にシーケンシングからデータの解析、検証まで一括して実践している。本研究においては、ゲノムワイドな解析の実験プロトコルとソフトウェアの両面から精度の向上に取り組んできた。次世代シーケンサーによる遺伝子診断の医療現場への普及を目指す際には、正確かつ効率的なシステムの構築が望まれるため、現在利用できる手段を比較し、現時点で最適と思われるパイプライン（組み合わせ）を提示した。研究期間の後期に、欧米で次世代シーケンサーの臨床応用に関するガイドライン等が公表されたが、当研究班で開発したパイプラインはこれらのガイドラインに記載される要件を満たしていると判断できた。

研究協力者

小崎健次郎（慶應義塾大学医学部臨床遺伝学センター）
清水厚志（岩手医科大学 いわて東北メディカル・メガバンク機構 生体情報解析部門）
鳥居千春（慶應義塾大学医学部臨床遺伝学センター）

A. 研究目的

3年間の研究期間に、次世代シーケンサーはもはや次世代の解析技術ではなくなり、現時点で診療応用が期待される医療技術となりつつある。一方で臨床の現場で次世代シーケンサーを診療に活用するためには、品質保証・医療経済・倫理的な配慮が必要である。これらの潜在的な問題点を念頭に置いて、診療上実用に供することの可能なシステムの構築を試みた。

B. 研究方法

① 標的遺伝子は3レベルに設定した。まず、特定領域の遺伝子群100個程度、全遺伝子、そして、現時点までに疾患との関与が明確になっている約5000個の遺伝子群である。標的遺伝子群について標準的な登録番号（カリフォルニア大学サンタクルツ校ゲノムデータベース UCSC ID）に変換した。患者ゲノムDNAからこれらの遺伝子に対応する領域を効率的に回収するためのオリゴヌクレオチドアレイを設計した。具体的にはAgilent社のカスタムマイクロアレイを用いることとし、同社のマイクロアレイ設計ソフトウェアである「eArray」を使用した。

特定の遺伝子群のエクソンを標的としたり、全遺伝子の全エクソンを対象とする場合、すなわち全エキソーム解析では、アジレント社の

SureSelectキットを使用した。一方で、ヒトの既知の疾患原因遺伝子のすべて（約5000遺伝子）を対象とする場合すなわちmedical exomeの解析の際には、イルミナ社のTruSightOneキットを用いた。

② ヒトゲノムDNAの前処理

市販ヒトDNAに対して、インターカレーション法による精密なDNA濃度測定を行い、Covaris社製超音波破碎機を用い、至適条件下で、約300bpへ断片化した。これをベックマン社AMPure磁性ビーズにより精製しSureSelectターゲットエンリッチメントライブラリーとハイブリダイズさせ、解析対象となる稀少遺伝性疾患に関連する遺伝子の近傍を濃縮した。コントロール（市販）のゲノムDNAはインデックスタグを付加し増幅し、イルミナ社製次世代シーケンサーMiSeqないしHiSeqを用いて塩基配列粗配列を得た。品質チェックを経て参照ゲノムDNA配列上へのマッピングを行った。

③ 変異の同定のためのデータの可視化

得られたDNA配列データをヒトの標準的なゲノム配列に整列（アラインメント）した。アラインメントには各種のプログラムを使用した。バリエーションに対して、各種のプログラムによりアノテーションとフィルタリングを行った。バリエーションの検出後のデータ解析にはグラフィ

カルユーザーインターフェースを有するプログラムを使用した。

④ DNA配列データのクオリティの確認

得られたDNA配列データがターゲットとした遺伝子領域を十分にカバーしているかを確認するため、FrommoltによるソフトウェアNGSrich (<http://sourceforge.net/projects/ngsrich/>) を用いて系統的に評価した。

解析にあたっては個人情報の保護に関する法律を踏まえ、文部科学省・厚生労働省・経済産業省「ヒトゲノム・遺伝子解析研究に関する倫理指針」、文部科学省・厚生労働省「疫学研究指針」を遵守し、倫理委員会の承認の下に研究を実施した。

C. 研究結果

① 解析系の設計

特定の標的領域を網羅的に解析する系を設計した。これら遺伝子の解析体制を整備した。次世代シーケンサーからの粗配列の出力ターゲット濃縮（アジレント社SureSelect）とデスクトップ型シーケンサー（イルミナ社MiSeq）を用いて対象疾患と関連する可能性のある遺伝子を網羅的に解析した。ヒト遺伝性疾患のうち、特に頻度が高い疾患を網羅的に解析する場合には、medical exome解析による方法が有用と思われた。

標的遺伝子解析については、NextGene等のソフトウェア会社の製品で十分に対応が可能と判断されたが、whole exomeやmedical exome解析については、大学等が開発して改訂・更新の周期の短いオープンソースのプログラム群を組み合わせて使用する方法を採用した。

② 粗配列のアラインメント

次世代シーケンサーからはDNA配列と塩基毎の精度がfastqフォーマットで出力される。そこで、得られたfastqファイルをヒト参照配列にオープンソフトウェアであるbwa (<http://bio-bwa.sourceforge.net/>) によりアラインメントした。さらに、GATK (IndelRealigner) によって局所的に再度のアラインメントを行い、アラインメントの質を向上させた。変異の検出には、GATK (UnifiedGenotyper) とイルミナ社が公開しているvariant caller、ベラー医科大学のatlas2 (米国ベラー大学)、コーネル大学のfreebayesなどを使用し、比較した。いずれもデフォルトの解析条件で分析を行ったところ、freebayesとatlas2は擬陽性率が高く、GATK unified genotyperとイルミナvariant callerを併用している。

③ 変異探索と原因候補遺伝子のリストアップ
アラインメントの結果を種々のソフトウェアにより分析し、アノテーションとフィルタリング

をおこなった。

Variant Effect Predictor (英国ウェルカム財団サンガー研究所とヨーロッパ生命情報研究所)、Snpeff (米国ウエイン大学)、annovar (フィラデルフィア小児病院)、alamut-ht等を用いた。アミノ酸置換の病的意義の解釈について香港大学 Li らが開発した、KggSeq (<http://statgenpro.psychiatry.hku.hk/limx/kggseq/>) を導入した。KggSeqでは10以上の予測プログラムを組み合わせた判定法である。Li MX, Gui HS, Kwan JS, Bao SY, Sham PC. A comprehensive framework for prioritizing variants in exome sequencing studies of Mendelian diseases. *Nucleic Acids Res.* 40:e53, 2012

Variant PredictorとSnpeffはバリエーションの国際標準フォーマットであるVCFファイルを入力ファイルとして、処理後、再びVCFフォーマットとして出力される点で、パイプラインの中間での使用が可能であった。一方で、annovarとalamut-htはVCFファイルの出力は不可能であり、パイプラインの末尾にのみ配置が可能であった。

alamut-htは、変異により潜在的なスプライスサイト（ドナーサイト・アクセプターサイトとも）を検出できる点で有用と考えられた。Alamutは年間ライセンス契約を要するサービスである点で、問題を残していた。代替プログラムとしては香港大学のKggSeqが有用と期待された。アミノ酸の非同義置換の病的意義の解釈については、KggSeqが、10種以上の予測プログラムの予測結果を組み合わせて、判定を行う点で優れていると考えられた。

④ DNA配列データのクオリティの確認

NGSrichによれば、アッセイ全体におけるカバーレージの統計的な情報値とともに、カバーレージが不十分であるエクソンを明確にすることが可能であった。ACMGのガイドラインにはカバーレージの状況についても、最終報告書に記載することが推奨されており、NGSrichの併用が有用であると考えられた。

D. 考察

我々が構築したエクソーム解析のシステムが、未知の疾患原因遺伝子の探索や非典型例の確定診断に有用であることが実証された。次世代シーケンサー解析の遂行には複数の過程が存在する。各過程について多くのプログラム群が国内外の大学等より発表されている。解析上の必要に応じて自由に組み合わせることができる点は、有用である一方、各プログラムの入出力を適切にインターフェースするためには一定のプログラミングのスキルを要するため、次世代シーケンサー解析の臨床現場への普及の妨げとなって

いるのが現状である。

大量の検体を同時並行的に処理しうるパイプラインを構築するとともに、医師が自身でインタラクティブにフィルターを掛けたり、シーケンサーからの生データを確認して、個別のバリエーションの病的意義を効率よく判定可能とする環境を医療の現場の中で構築してゆくかが今後の課題と思われる。

研究期間の最終年度に、アメリカ臨床遺伝専門医会 ACMG から、ACMG clinical laboratory standards for next generation sequencing が発表された。本研究班で最適化を行ってきた要点と概ね一致していた。今後は、系の再現性を担保する方策の確立につとめるとともに、当該ガイドラインに示される報告書のテンプレート等を参考に、依頼医師に対して間違いのない結果の伝達に留意したい。上記ガイドラインによれば、電子カルテへの記載法については、electronic health record 国際基準案 (HL7; <http://www.hl7.org>) に従うことが奨励されている。われわれのパイプラインの末尾には HL7 形式にデータをリフォーマットするスクリプトを加えた。国内標準形式となるように関係者への普及を図りたい。

E. 結論

正確かつ効率的なシステムの構築を目指して、現在利用できる手段を比較し、現時点で最適と思われるパイプライン（組み合わせ）を提示した。研究期間の後期に、欧米で次世代シーケンサーの臨床応用に関するガイドライン等が公表されたが、当研究班で開発したパイプラインはこれらのガイドラインに記載される要件を満たしていると判断できた。

F. 研究発表

1. 論文発表

1) Sasaki, T., Niizeki, H., Shimizu, A., Shiohama, A., Hirakiyama, A., Okuyama, T., Seki, A., Kabashima, K., Otsuka, A., Ishiko, A., Tanese, K., Miyakawa, S., Sakabe, J., Kuwahara, M., Amagai, M., Okano, H., Suematsu, M., Kudoh, J. Identification of mutations in the prostaglandin transporter gene *SLCO2A1* and its phenotype-genotype correlation in Japanese patients with pachydermoperiostosis. *J. Dermatol. Sci.*, 68(1):36-44 (2012).

2) Miyamoto K, Suzuki N, Sakai K, Asakawa S, Okazaki T, Kudoh J, Ikeno M, Shimizu N. A novel mouse model for Down syndrome that harbor a single copy of human artificial chromosome (HAC) carrying a limited number of genes from

human chromosome 21. *Transgenic Res.*23(2);317-29(2014).

3) Kubo A, Shiohama A, Sasaki T, Nakabayashi K, Kawasaki H, Atsugi T, Sato S, Shimizu A, Mikami S, Tanizaki H, Uchiyama M, Maeda T, Ito T, Sakabe J, Heike T, Okuyama T, Kosaki R, Kosaki K, Kudoh J, Hata K, Umezawa A, Tokura Y, Ishiko A, Niizeki H, Kabashima K, Mitsuhashi Y, Amagai M. Mutations in *SERPINB7*, Encoding a Member of the Serine Protease Inhibitor Superfamily, Cause Nagashima-type Palmoplantar Keratosis. *Am J Hum Genet.* 93(5):945-956(2013).

4) Mutai H, Suzuki N, Shimizu A, Torii C, Namba K, Morimoto N, Kudoh J, Kaga K, Kosaki K, Matsunaga T. Diverse spectrum of rare deafness genes underlies early-childhood hearing loss in Japanese patients: A cross-sectional, multi-center next-generation sequencing study. *Orphanet J. Rare Dis.*8(1);172(2013).

5) Niizeki H, Shiohama A, Sasaki T, Seki A, Kabashima K, Otsuka A, Takeshita M, Hirakiyama A, Okuyama T, Tanese K, Ishiko A, Amagai M, Kudoh J. The novel *SLCO2A1* heterozygous missense mutation p.E427K and nonsense mutation p.R603* in a female patient with pachydermoperiostosis with an atypical phenotype. *Br J Dermatol.* [Epub ahead of print] (2013).

6) Kanda S, Sasaki T, Shiohama A, Nishifuji K, Amagai M, Iwasaki T, Kudoh J. Characterization of canine filaggrin: gene structure and protein expression in dog skin. *Vet Dermatol.*24(1);25-31(2013).

7) Sasaki T, Shiohama A, Kubo A, Kawasaki H, Ishida-Yamamoto A, Yamada T, Hachiya T, Shimizu A, Okano H, Kudoh J, Amagai M. A homozygous nonsense mutation in the gene for *Tmem79*, a component for the lamellar granule secretory system, produces spontaneous eczema in an experimental model of atopic dermatitis. *J Allergy Clin Immunol.*, 132(5):1111-1120 (2013).

2. 学会発表

1) 工藤 純、ゲノム解析による遺伝性疾患の原因解明、第16回九州基礎皮膚科研究会 特別講演、福岡、2012年11月17日

2) 佐々木貴史、新関寛徳、清水厚志、塩濱愛子、開山麻美、奥山虎之、関 敦仁、椛島健治、大塚篤司、石河晃、宮川 俊一、天谷雅行、岡野栄之、末松誠、工藤 純、次世代シーケンサーを用いたエキソーム解析による肥厚性皮膚骨膜症原因遺伝子 *SLCO2A1* の同定、日本人類遺伝学会第57回大会、京王プラザホテル、平成24年10月25日～27日

3) Kudoh J, Sasaki T, Shimizu A, Shiohama A, Hirakiyama A, Okuyama T, Seki A, Kabashima K, Otsuka A, Ishiko A, Tanese K, Miyakawa S, Sakabe J, Kuwahara M, Amagai M, Okano H, Suematsu M, Niizeki H. Identification of mutations in the prostaglandin transporter gene *SLCO2A1* in Japanese patients with pachydermoperiostosis. American Society of Human Genetics 2012 Annual Meeting, San Francisco, November 6-10, 2012.

G. 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

厚生労働科学研究費補助金
難病・がん等の疾患分野の医療の実用化研究事業（難病関連分野）
総合研究報告書

疾患の遺伝要因解明のための多層オミックス統合解析の基盤構築

研究分担者 清水厚志
岩手医科大学 いわて東北メディカル・メガバンク機構 生体情報解析部門
部門長代理 特命教授

研究要旨

次世代シーケンサー（NGS）の技術革新と解析技術の発展を受け、比較的安価にヒトのゲノム、オミックス情報を入手することが可能となった。大規模なヒト多型データの収集も世界規模で進められており、疾患原因遺伝子の同定も促進された。しかし、現時点においても新規の原因遺伝子同定率は3割程度にとどまる。この原因の一つとして、アミノ酸置換によるタンパク質機能異常の他に、遺伝子制御領域での塩基置換やDNAメチル化の異常による遺伝子発現量の変化が疾患発症に関与すると考えられる。そこで、本研究では、1) 患者ゲノム配列を解析することで疾患原因の同定するための基盤構築、2) 遺伝子発現量を解析することで疾患原因の同定するための基盤構築、3) DNAメチル化状態を解析することで疾患原因の同定するための基盤構築、の3つを目的とし、遺伝性疾患を多層オミックスの観点から解明する次世代型遺伝子解析技術の基盤構築を進めた。

研究協力者

八谷剛史^{1,2}、大桃秀樹²、志波 優²、古川亮平²

1) いわて東北メディカル・メガバンク機構 生体情報解析部門

2) いわて東北メディカル・メガバンク機構 メガバンクデータ管理部門

A. 研究目的

近年次世代シーケンサーの台頭により多数の遺伝性疾患の原因遺伝子が同定されている。しかし、既知の遺伝性疾患であるにも関わらず、変異が未だ同定できていない患者が少なからず存在する。その理由の1つとして同一疾患で原因遺伝子が異なるlocus heterogeneityなどが考えられる。

そこで、数十人から数百人の特定の遺伝性疾患患者のゲノムDNAから当該疾患に関与すると疑われる数百程度の限定した遺伝子のみについて、次世代シーケンシング法により配列情報を入手し、既知の一塩基多型（SNPs）との比較により患者特有の変異／多型を抽出することで、各疾患の原因となる特徴的な遺伝子変異をカタログ化することを第一の目的とした。

続いて、疾患原因遺伝子候補の同定のため、あるいは原因遺伝子はほぼ確定しているが、コーディング領域に変異が見つからない患者において、RNAを直接解析対象として用いるために遺伝子発現解析の基盤技術を構築することを第二の目的とした。

RNAを解析することは今後の変異解析に必須の技術であるが、患者特異的iPSではなく、患者

血液を検体として用いる場合はRNAの不安定性が問題となる。そこで、DNAメチル化解析技術と組み合わせることで、疾患原因遺伝子同定の精度の向上ならびに遺伝子制御領域での変異の影響を検証するため、DNAメチル化解析基盤の構築を第三の目的とした。

さらに、得られた変異の検証方法として、小型魚類であるメダカを用いた原因候補遺伝子の検証法の確立を目的とした。

B. 研究方法

1. Agilent SureSelect Custom Target Enrichment kit の設計

各分担グループから過去の臨床報告や、分子レベルでの相互作用から推定して提示された150個から300個程度の遺伝子をUCSC gene IDに変換し、Agilent社のカスタムマイクロアレイ設計ツールである「eArray」により目的遺伝子断片を効率よく回収するCustom Target Enrichment kit（以下カスタムキット）を設計した。eArrayで自動設計したカスタムキットにはGC含量が高い領域や反復配列の混入を避けるために、複数のエクソン領域でプローブが未設計であった。そこで、プローブの配列とUCSC

のエクソン位置情報を比較し、可能な限り未設計領域へ新たにプローブを追加設計し、Agilent社に合成を依頼した。

2. Illumina社製次世代シーケンサーMiSeqによる解析

デスクトップ型次世代シーケンサーであるMiSeqは小型ではあるが、出力量は2Gbにもおよび、数年前のフラッグシップ機と同等の性能を持つ。そこで上記カスタムキットと、既存のキナーゼのみをターゲットとしたSureSelect Human Kinome Kit（以下キノームキット）を使用しMiSeqにより患者ゲノム配列の解読を行った。

3. 疾患解析パイプラインの構築

主にアカデミックフリーのソフトウェアを利用して、

- i) NGS配列データ (fastq) から校正済みゲノムアライメントファイル (bam) の作成
- ii) ヒト参照配列との比較による多型の抽出
- iii) 得られた個人の多型情報と公開多型データベース (dbSNPs、1KG) の比較による変異候補遺伝子の抽出

を行うシステムを構築した(清水 厚志、細胞工学. 2011)。しかし、この手法によっても個人ごとに300ヶ所以上のレア多型/変異が検出され、単一遺伝子疾患と考えられる患者1名分のデータから原因遺伝子を同定することが不可能な症例が多数あった。一方で、公開多型データベースの登録情報すべてをフィルターとして使用すると、数千人に1アレルというごく稀な多型/変異まで疾患候補変異から除外してしまうため、公開多型データの健常人ボランティアの中に偶然患者と同じ変異を持っていた場合に擬陰性が生じる可能性があった。また、頻度の高い一塩基多型 (SNPs) に隣接して変異が生じた場合に、既存の手法ではそれぞれを独立の変異として処理してしまい、異なるアレル上の正確なアミノ酸置換を再現できない問題もあった。

そこで、

- i) 複数の公開多型データベースを統合した最新の既知多型リストへの更新
- ii) 頻度情報を考慮したフィルターの作成
- iii) 連続してレファレンスと異なる場合の同一アレル内多型同定機能の追加
- iv) 複数の検体間の共通変異抽出機能の追加を行った。

また、コマンドライン入力に慣れていない医師、研究者が解析研究を行う状況を考慮し、市販ソ

フトウェアでは評価の高いCLC Genomics workbench (CLC) と我々が構築した解析システムとの比較検討を行った。

4. 発現変動遺伝子同定システムの構築

ゲノムDNA解析によって同定された変異のうち、ナンセンス変異やエクソン-イントロンジャンクション等の変異を疾患の原因と推定することは比較的容易である。一方で、イントロンやプロモーター等の制御領域の変異はゲノムDNA解析ではその影響を推定することがほぼ不可能である。そこで、本事業分担研究者の赤松らにより、将来的に疾患患者特異的iPS細胞の樹立が容易となるという予測のもと、分化誘導させた患者特異的細胞とコントロール細胞の遺伝子発現量を次世代シーケンサーのデータから比較できる解析システムを構築した。

5. メチル化変動遺伝子同定システムの構築

複数の健常人ボランティアからBD バキューティナ® CPT™採血管で採取した血液を匿名化後、分離した末梢血単核球細胞 (PBMC) からDNAを抽出した。匿名化された1個人のDNA 50ngをDNA Methylation-Gold Kit, EZ (Zymo Research社) によりバイサルファイト処理を行い、EpiGnome Methyl - Seq Kit (epicentre社) によりライブラリーを調整した。調整したライブラリーはバイサルファイト処理によりATリッチに傾いてしまうため、シーケンシングの際の塩基バランスを整えるためにコントロールであるphiX DNAを30%量混合させ、HiSeq2500のRapid mode, single read 全ゲノムバイサルファイトシーケンシング (WGBS) を行った。WGBS後にHiSeq2500から得られたbclファイルを昨年度までに構築した前処理プロセスにより、高品質なfastqファイルに変換した。WGBSは多型解析やトランスクリプトーム解析と異なり、ゴールドスタンダードとなりうるパイプラインが確立されていないため、複数のソフトウェアで検討を行った。

(倫理面への配慮)

上記ヒト検体を用いた研究は文部科学省・厚生労働省・経済産業省「ヒトゲノム・遺伝子解析研究に関する倫理指針」に準拠した研究計画に基づき、慶應義塾大学医学部倫理委員会に承認を受け実施した。

6. モルフォリノアンチセンスオリゴ (MO) による疾患候補遺伝子のノックダウン

次世代シーケンシング法により得られた疾

患原因候補遺伝子の一つをモデルとして、メダカ胚でのノックダウン解析を試みた。Ensembl (<http://www.ensembl.org/>)にて、ヒト遺伝子名からメダカオルソログ遺伝子を検索して配列情報を入手した。ESTデータベースと比較ゲノム解析により遺伝子構造モデルを推定した。予想cDNA配列の開始コドンに対応するMOを作製した。作製したMOを一細胞期のメダカ受精卵にインジェクションし、経時的に表現型の変化を観察した。

(倫理面への配慮)

本動物実験は慶應義塾大学動物実験委員会のガイドラインに従った。また、本研究に関する遺伝子組換え、動物実験はすべて適切な委員会の承認を受けた。

C. 研究結果

1. MiSeqによる次世代シーケンシング

市販のキノームキット (約 600 遺伝子対象) と MiSeq による患者 1 名のゲノムシーケンシングでは、対象領域当り平均 600 本の配列を得ることができた。一方、カスタムキット (約 300 遺伝子対象) と MiSeq では患者 4 名で対象領域当り平均 200 本の配列を得ることができた。

2. 変異探索システムの改良

2.1. 既知多型データベースからの多型情報の入手と既知多型リストの更新

dbSNP137、ESP6500、1KG low coverage、1KG 日本人エクソームの多型データを統合し、既知多型リストを更新した。特に日本人エクソームのデータに関しては我々が構築した多型抽出システムを使用して、同じ処理/同じ精度の多型リストを独自に作成した。その結果、昨年度の既知多型リスト (48, 455, 489ヶ所) におよそ400 万ヶ所を追加できた (52, 659, 387ヶ所)。

2.2. 頻度情報を考慮したフィルターの作成

それぞれの公開多型データに付随した頻度情報 (MAF; Minor allele frequency) から $MAF > 0.01$ (>1%) の多型のみを抽出した高頻度多型フィルターを作成し、効果を検証した (表 1 : 後述)。

2.3. 連続塩基置換のアレル同定

構築した変異探索システムには GATK (Genome Analysis tool kit; <http://www.broadinstitute.org/gatk/>) を利用しているが、GATK で抽出された塩基置換リスト

にはアレル情報が欠失している。そこで、抽出された塩基置換がゲノム位置上で連続していた場合に、構成済みアライメントデータに含まれる配列情報から、それらの塩基が同一アレル由来か判定するシステムを構築した。HapMap の検体 (NA18943) を用いて解析したところ、CACNA1H のコーディングエクソン上に同一アレル由来の連続塩基置換を同定した (図 1)。

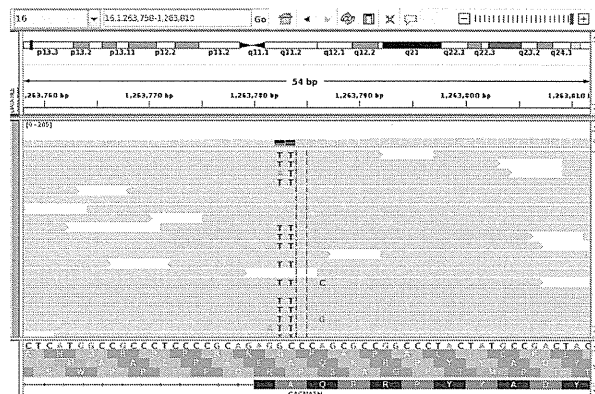


図 1 : NA18943 で同定された連続塩基置換

2.4. 複数の検体間での共通変異の抽出

変異探索システムで候補として上がったゲノム上の変異から対応する遺伝子のリストを作成して、患者内でのみ変異が見つかった遺伝子を抽出する機能を追加した。

2.5. 変異探索システムの検証

2.4. までの追加機能の効果を検定するため、すでに原因遺伝子が同定されている常染色体劣性遺伝子疾患患者エクソーム解析データで検証を行った (表 1)。

表 1 : 原因遺伝子の候補リスト中の順位*

A:	B:	C:	D:
完全新規	MAF < 0.01	MAF < 0.01	MAF < 0.01
dbSNP137	dbSNP137	dbSNP137	dbSNP137
JPT Exome	JPT Exome	JPT Exome	
1KG	1KG		
ESP6500	ESP6500		
2位	6位	10位	131位

*HLA除く

現在構築したシステムは擬陰性をできるだけ排除するために、シーケンス精度が比較的低い塩基由来の候補部位も残すフィルターに設定している ($QV \geq 20$, $Depth \geq 4$)。そこで、IGV (Integrative Genomics Viewer) を用い、目視にて表 1 A-C の上位遺伝子群の候補部位を確認すると、候補遺伝子の大多数はシーケンス精度が今回設定した精度の下限値であった。また、原因遺伝子以外の十分なシーケンス精度持つ

変異候補は異なる家系であるにもかかわらず同一の塩基置換であることがわかった。

2.6. 市販ソフトウェアとの比較

2.5. の検証で我々のシステムにより効果的に原因遺伝子を同定できることが確認できた。しかし、本システムはLinux上で動くコマンドラインベースで構築されているため、一定レベルのコンピュータスキルを必要とする。そこで、コマンドライン入力に慣れていない医師、研究者が市販のソフトウェア (CLC) を使用した際にどの程度の精度で変異を同定することができるか、主にアライメントの機能について比較検討した。同一個人のエクソーム解析の配列データを本システム、CLC、bowtie2でレファレンス配列にマッピングし、特にアライメントの結果が大きく異なると予想された繰り返し配列を含むindel (insertion - deletion) の領域をアライメントデータから抽出した。本システムではindel周辺も正確にアライメントできていた (図2上段) のに対し、CLCでは大半が考慮されずにトリミングされており (図2中段)、比較のため使用したbowtie2ではアライメントが正確には行われず、indelではなく多数のSNPsとして同定された (図2下段)。

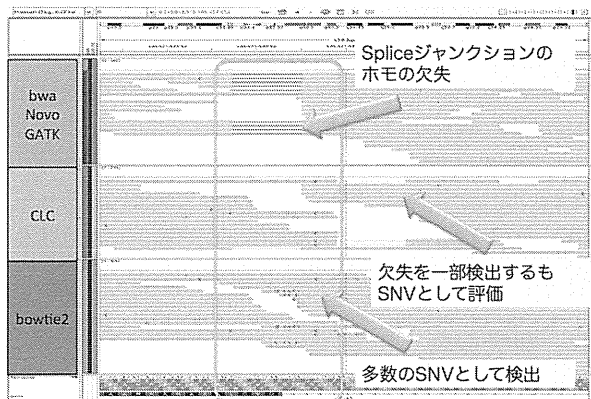


図2：複数のアライメントソフトの比較

3. 発現変動遺伝子同定システムの構築

現時点で疾患患者特異的iPS細胞のデータを入手していないため、Human Body Map 2.0 からヒト組織別RNA-seqの配列データ

(<http://www.ebi.ac.uk/arrayexpress/experiments/E-MTAB-513/>) を入手し、システム構築のモデルデータとして、以下のシステムを構築した。

- i) 得られた配列データに含まれるアダプター配列とクオリティ値が低い塩基の除去
- ii) ヒトレファレンス配列への遺伝子構造を考慮したマッピング

- iii) 遺伝子毎の発現量の推定
 - iv) 得られた予測発現量の差分処理
 - v) 組織間で発現量が有意に異なる遺伝子群の抽出と描画
 - vi) Gene Ontology 抽出
 - vii) 既知の遺伝子ネットワーク上への描画
- 差分処理iv)までをLinux上で処理し、その後の処理をPC上のR packageで処理することで効率的にネットワーク解析まで行うことができた。

4. メチル化変動遺伝子同定システムの構築

マッピングソフトの最適化のため、DNAメチル解析で広く使われているbismark、smoothingが可能なパッケージとの連携が可能なBsmooth、市販ソフトであり多型解析で実績のある novoalign の3者を比較した。bismarkを用いたところ、マップ率は55%程度であった。Bsmoothでは91%、novoalignでは当初89%のマップ率であったが、novoalignのパラメーターを調整し、最終的に92%のマップ率まで向上させることができた (図3)。

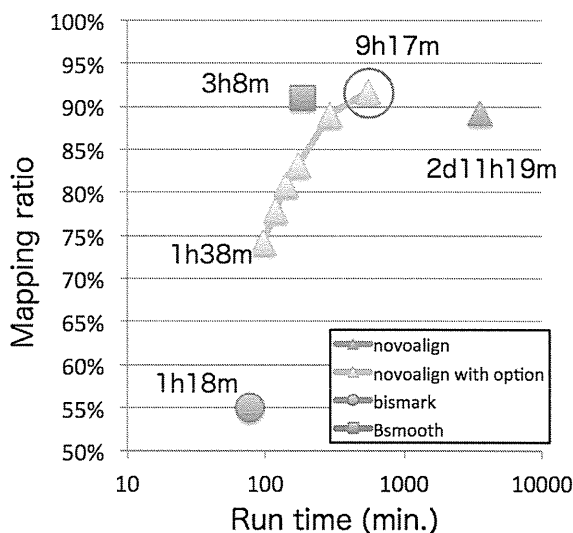


図3：複数のアライメントソフトの比較

続いて、メチル解析ソフトについて検討した。methylKitは複数の検体間でのdifferentially methylated region (DMR)やdifferentially methylated cytosine (DMC)の検出、相関解析や階層的クラスタリング解析、主成分分析が可能なRパッケージである。bismarkから出力されるbamファイルをシームレスにDNAメチル解析できることが特徴の一つであるが、WGBSの様な大きなデータだと、読み込みに非常に時間がかかり、特に複数検体の取り込みでは顕著である。そこで、methylKitに読み込まれたbamファイル

のR内部での変換後のフォーマットを確認し、予め、解析サーバ上でbamファイルを整形してメチル化部位のアノテーション情報に変換することで、methylKitでの取り込みと解析する時間を劇的に短縮することができた。また、novoalignによるマッピングから得られたbamファイルは、そのままではmethylKitで読み込むことができないが、samtoolsでWatson鎖とCrick鎖に分離し、それぞれmpileupを用いてメチル化部位候補をすべて抽出した後にin-house プログラムで整形することで、同様にmethylKitで解析することが可能となった(図4, 5)。

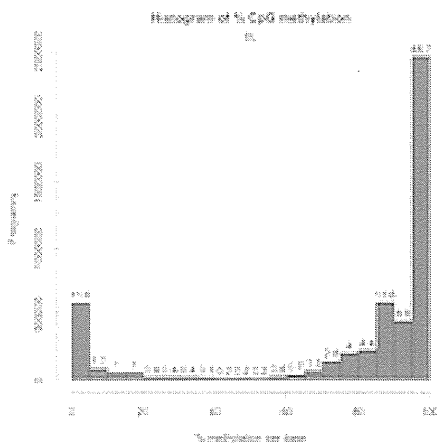


図4：メチル化CpG部位の頻度

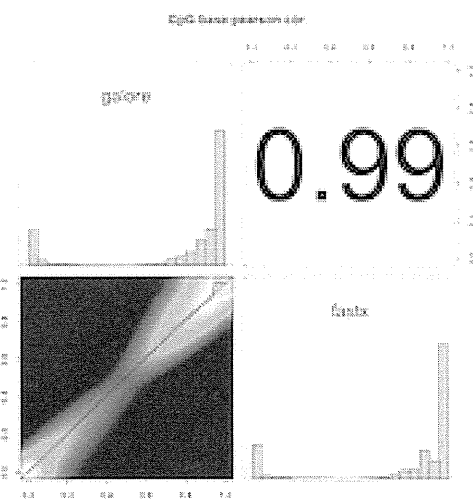


図5：methylKitによるトリミングソフトの比較

5. メダカ胚を用いたノックダウン

開始コドンに設計したMOとコントロールMOの間での明瞭な表現型の差を観察する事はできなかった。

D. 考察

単一遺伝性疾患と考えられるが原因遺伝子が未知の症例において、エクソーム解析は一定

の成果をあげたが、現時点においても原因遺伝子が同定できない例も多く、その理由の一つとして、遺伝子発現量変化が挙げられる。同様にヒトゲノム計画後に急激に発展した

genome-wide association study (GWAS)により、13,000を超える疾患感受性多型が発見されているが、その80%以上は非コード領域に位置するため、疾患に影響のある遺伝子が同定できていない。ENCODEプロジェクトの結果からもnon-coding領域の大部分がいずれかの転写に関わると推察されており、今後の疾患解析にはコーディング領域だけではなく、遺伝子制御領域の多型/変異の同定と意味付け(アノテーション)が必須である。本研究はゲノム多型、遺伝子発現、DNAメチル化の3つの基盤構築を進めてきたが、今後、これらの基盤技術と、疾患患者特異的iPSとの連携により、効率的な新規原因遺伝子同定のシステム構築が可能となる。

E. 結論

3年間の本プロジェクト推進の中で、初年度に構築した多型解析からの疾患原因遺伝子の同定システムはエクソームおよびターゲットリシーケンシングの双方で成果をあげることができた(初年度、二年度業績、および本年度業績1、2)。また、二年度に開発した発現解析システムについても、iPSと分化誘導した細胞との変動遺伝子の抽出や、組織特異的発現遺伝子の同定が可能であることを確認した(論文未発表)。本年度に構築したシステムについても患者特異的なDNAメチル化を同定することが可能であり、今後、これらゲノム・オミックス解析基盤を活用した多層オミックス解析による疾患原因遺伝子の同定を推進する基盤の構築を達成できた。

F. 研究発表

1. 論文発表

- 1) 清水 厚志*, 佐々木貴史：基礎の基礎. 細胞工学. 2011, 30:790-795 *本特集号「次世代シーケンサーを使いこなす」監修
- 2) 清水 厚志：ヒト・マウスの全エキソンリシーケンシングと疾患原因遺伝子の同定. 細胞工学. 2011, 30:808-814
- 3) 清水 厚志：次世代シーケンサーによるトランスクリプトーム解析. 臨床検査. 2011, 55:841-846
- 4) 清水 厚志：次世代シーケンサーによる発現変動遺伝子の同定. Medical Science Digest. 2012; 38: 587-588
- 5) Mutai H, Suzuki N, Shimizu A, Torii C, Namba K, Morimoto N, Kudoh J, Kaga K, Kosaki K and Matsunaga T. Diverse

spectrum of rare deafness genes underlies early-childhood hearing loss in Japanese patients: A cross-sectional, multi-center next-generation sequencing study. *Orphanet Journal of Rare Diseases*. 2013; 8:172

- 6) Takenouchi T, Shimizu A, Torii C, Kosaki R, Takahashi T, Saya H and Kosaki K. Multiple Café au Lait Spots in Familial Patients With MAP2K2 Mutation. *American Journal of Medical Genetics*. 2013; 164A: 392-396.
- 7) 清水 厚志: 次世代シーケンサー入門. 遺伝子医学 MOOK 別冊「いまさら聞けない『遺伝医学』」. 2014, in press.
- 8) 清水 厚志, 八谷 剛史, 田原 康玄: ヒトゲノム・オミックス情報のコホート研究への応用. 実験医学別冊「次世代シーケンス解析スタンダード」. 2014, in press.

2. 学会発表

- 1) Shimizu, A, Minoshima S. and Shimizu, N.; Retrotransposition of core duplication triggered formation of multiple low copy repeats (LCRs) during primate evolution. 12th International Congress of Human Genetics / 61th The American Society of Human Genetics Annual Meeting, Montreal, Canada (Oct. 2011) 10/11-15
- 2) 清水 厚志, 古家育子, 佐々木貴史, 工藤純: 慶應義塾大学医学部内での受託解析システムの構築. 第2回次世代シーケンサー現場の会大会. 2012 5/26 大阪.
- 3) 清水 厚志: Miseq による既知疾患遺伝子群のターゲットリシーケンシング. 2012 イルミナ次世代シーケンサーユーザーフォーラム. 2012 6/20 東京 (招待講演)
- 4) 清水 厚志: 比べてみよう CLC Genomics Workbench とオープンソース; 市販ソフトとアカデミックフリーソフトの共存. CLCbio ユーザーミーティング 2012. 2012

6/26 東京 (招待講演)

- 5) 鳥居 千春, 丸岡 亮, 清水 厚志, 小崎里華, 小崎 健次郎: 次世代シーケンサーを用いた先天奇形症候群の網羅的診断. 第52回日本先天異常学会学術集会. 2012 7/16-18 東京.
- 6) 清水 厚志: ゲノムプロジェクトからパーソナルゲノムまでの10年. 次世代シーケンシングが切り開く医学研究と展望 京都大学 iPS 細胞研究所ゲノム・エピゲノム解析コアファシリティ. 2012 7/31 京都 (招待講演).
- 7) Shimizu A, Torii C, Suzuki N, Mutai H, Kudoh J, Kosaki R, Matsunaga T, Kosaki K. Rapid and efficient mutation detection in the hundreds of target genes by bench-top next generation sequencer with custom target capture method. The American Society of Human Genetics 62th Annual Meeting, San Francisco, USA (Nov. 2012) 11/6-11
- 8) 鈴木 直大, 務台 英樹, 鳥居 千春, 清水 厚志, 宮 冬樹, 難波 一徳, 工藤 純, 小崎 健次郎, 松永 達雄: カスタムターゲットリシーケンスによる難聴関連遺伝子の変異探索. 日本人類遺伝学会. 第57回大会, 2012 11/20-23 東京.
- 9) 清水 厚志. 遺伝性疾患患者ゲノムの Exome および Targeted resequencing 解析の実例. 第35回日本分子生物学会年会 アジレントテクノロジーランチョンセミナー. 2012 12/13 福岡 (招待講演).

G. 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

厚生労働科学研究費補助金
難病・がん等の疾患分野の医療の実用化研究事業（難病関連分野）
総合研究報告書

次世代シーケンサーを用いたバイオインフォマティクス解析による
先天性異常症の遺伝要因の解明に関する研究

研究分担者 宮 冬樹

独立行政法人理化学研究所 横浜研究所 統合生命医科学研究センター
医科学数理研究グループ・リサーチアソシエイト

研究要旨

次世代シーケンスデータから疾患原因変異を同定するためには、膨大なシーケンスビッグデータからできる限り見逃し（偽陰性）を無くし、かつエラー（偽陽性）も少なくし、究極的には1つの原因変異に絞り込む必要がある。我々はそのために独自の解析パイプラインを構築し、実際に偽陰性・偽陽性率のいずれも極めて0に近いことを実験にて確認し、その疾患変異探索システムを完成させた。当解析パイプラインを用いて約300検体の先天性疾患検体のexonシーケンス解析を実施し、新規・既知を合わせ数十家系の疾患原因変異を同定した。このシステムは一般的な1塩基変異（SNV）のみならず挿入・欠失（indel）も高精度で検出でき、当初の予想以上に複合ヘテロで発症にいたるケースが複数例同定され、またSNVとindelの複合ヘテロも見られ、製作した解析システムの精度が確認できたと共に、今後を含めた疾患解析時における重要な着目点を見出した。新規の原因遺伝子変異候補が複数発見され、基礎科学的な側面の有用性が確認されたと同時に、確定診断に至っていない患者において既知の原因遺伝子変異が認められたことにより確定診断に至った例もあり、臨床的な側面での実用性も確認された。解析プログラムの最適化による高速化を施し、基礎研究のみならず迅速さが求められる臨床検査においても対応しうるシステムを構築した。

A. 研究目的

ヒト疾患についてはこれまで多数の原因遺伝子が明らかになっており、2013年の段階で疾患に関与している遺伝子が約6,000遺伝子、その遺伝子が関わる何らかの疾患関連の表現型として約13,300が知られているものの、未だに原因遺伝子が不明な先天性疾患も多く、また原因遺伝子が一部明らかになっている疾患についても同じ疾患であるが原因遺伝子が複数存在する場合も多く、変異の頻度が高い原因遺伝子以外については検査が難しく、治療法もほとんど存在しない場合が多かった。また、我々の研究班内で対象にした先天性疾患の中でも特に神経系疾患の場合は表現型が複数に跨って併発する場合も多く、原因遺伝子が不明なことにより確定診断が難しい場合も多くあった。近年、次世代シーケンサーという技術的なブレイクスルーにより、多数（ほぼ全て）の遺伝子について一度に配列解析が可能となり、低頻度あるいは新規の原因遺伝子探索が網羅的に可能となった。この技術を用いて、我々の研究班では従来の診断では同定できなかった先天性疾患の原因遺伝子同定を目指すことを目的としたが、そこには精度と検出力のさらなる強固さが

必要不可欠であった。そのために、我々は次世代シーケンス由来のビッグデータから新規の疾患原因遺伝子を特定するための解析パイプラインの構築を行い、その精度検証を実施しつつ最適化を試みた。完成した解析パイプラインを用いて研究班コンソーシアムから提供された数百例の先天性疾患検体のexome解析を実施し、各疾患患者の原因遺伝子探索を実施した。

B. 研究方法

遺伝子のexome解析の手法としては、主に2種実施し、その内訳は、類似の先天性疾患ごと（神経系疾患群あるいは難聴および奇形症候群）において既知または関連性が疑われる遺伝子セットを選定し、SureSelectカスタムターゲットエンリッチメントシステムで作製した特定遺伝子（284遺伝子または309遺伝子）のカスタムプローブを用いたexomeが一つ、もう一つはSureSelectXT Human All Exon V4/V5を用いた全遺伝子exomeである。検体が患者のみあるいは特定のカスタム遺伝子セット内の遺伝子に原因遺伝子が含まれることが予想される患者検体に関してはカスタムexomeを優先的に実施し、疾患家系検体が少なく

とも両親および発症患者の計3名以上が用意できる検体およびカスタムexomeでは原因遺伝子が検出されなかった検体に関しては全exomeを実施した。

実験の手法としては、Agilent社のSureSelect Target Enrichment Systemのプロトコルに準拠したが、端的に記載すると、患者由来の血液からDNAを抽出し、3ug分のDNAをCovaris社のDNA Shearing system M220にて150-200bpに断片化後、特定遺伝子のカスタムプローブまたは全遺伝子のプローブにて特定領域のゲノムDNAをキャプチャーし、次世代シーケンス用アダプターを付加させたのちDNAを増幅し、Illumina社の次世代シーケンサーGAIIまたはHiSeq2000にてペアエンドシーケンスを行った。データ量はカスタムexomeが2Gb程度、全遺伝子exomeが7Gb程度になるようにシーケンスを実施し、その平均読み深度 (depth, 同じ場所が何度シーケンスされたか) は最低50以上となるようにした。

解析のパイプラインの概要を図1に示した。次世代シーケンサーから得られたデータから、まずはPCRエラー等を含む可能性のある同一重複リード (duplication read) の除去を行い、次にBWAソフトウェアにてヒトリファレンスゲノムにマッピングを行い、リファレンスゲノムにユニークにマッピングされ、かつペアエンドのリードが互いに平均ペアエンドの距離 ± 2 SD (標準偏差) 以内に存在するリードのみを残した。その後、Samtoolsソフトウェアおよび独自に開発したプログラムを用いつつ、各種base qualityやmapping quality score、consensus score、SNP quality scoreなどのパラメータを調整し1塩基多様性(SNV; single nucleotide variation)を検出した。挿入・欠失 (indel; insertion and deletion) の検出の場合には、GATKソフトウェアを用いて、SNVが検出されたリードを含む配列をリアライメントし直し検出することで精度を上げた。次世代シーケンスにて読まれたある同一箇所のアレルの頻度が偏っていてもheterozygoteとコールされる場合があり、そのような頻度が偏っている場所はエラー率が高いことが明らかとなっていることから、同一箇所でも25%以上リファレンスと異なる塩基がコールされる場合のみ変異として出力するようにした。次世代シーケンスデータ

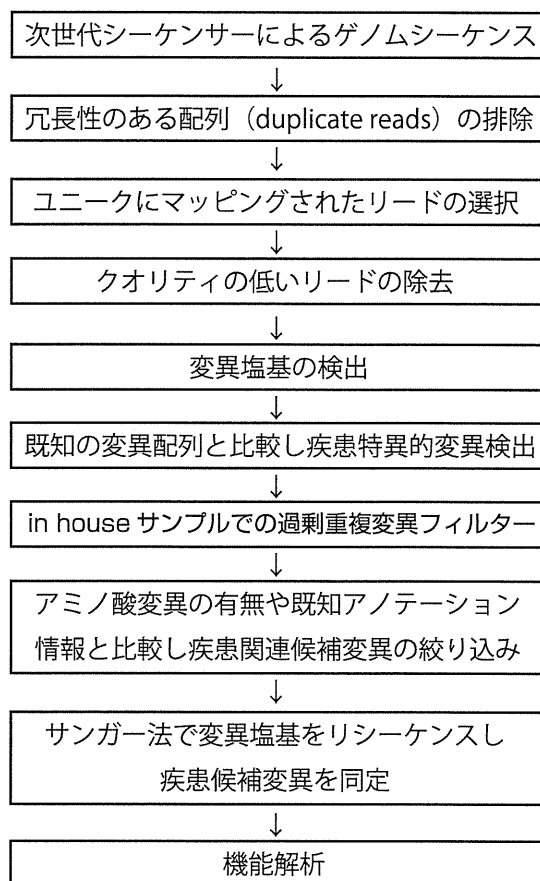


図1. 解析パイプラインの概要

から偽陽性ではない真の変異を抽出した場合でも、一般集団で有するcommon variantであれば先天性疾患の原因遺伝子である可能性は極めて低く、レアな変異のみに絞り込むことは極めて重要である。そこで既知の頻度の高めの多型 (SNP) を除くために、既知変異のフィルターとしてdbSNP、1000 Genomes Projectのデータ、ESP6500という3種のデータベースに登録されている多型情報 (約4,700万多型塩基) を利用した。一般的なフィルターでは変異が登録されている染色体上の位置情報のみからフィルターにかける場合も多いが、疾患要因変異と登録変異は同じ場所でも塩基が異なる可能性も十分あることから、我々のフィルターでは塩基の一致についても調査し、データベースに染色体上の同一位置で変異が登録されていても塩基が異なる場合はフィルターで除外しないようにした。また、一般集団でもわずかながら発症はしなくとも原因遺伝子変異の保因者である可能性もあることから、フィルターを2種類作製し、1つは既報のものを全て除く「Filter set 1」、もう一つはアレル頻度0.001を閾値として一般集団でも有する変異を残す緩めの「Filter set 2」として作製した。さらに既知の疾患関連変異として知られている変異で、一般集団におけるアレル頻度が0.01未満の変異はFilterにおいて除外せずに「Filter set 2」において残すようにし

た。これは、先天性疾患のみならず、他の疾患の原因として知られている変異が、実は先天性疾患の表現型とリンクしている可能性を考慮したことによる。既知疾患変異のデータベースとしては世界最大のヒトメンデル病疾患データベースであるHGMD Professionalと、NCBIのOMIMデータベース、dbSNPの登録変異のうち疾患との関与(示唆)が報告されているpathogenic SNPの3種を用い、合計で約8万塩基を疾患関連変異としてフィルターに用いた。次に、データベース上の既知のSNVにはないが、我々の実験系で読まれたデータ中で多数の検体が染色体上の同一箇所の変異を保有していた場合、その変異はプラットフォーム依存的な偽陽性である可能性、あるいは日本人特有の人種特異的な多型である可能性が大きいいため、我々の検体データ (*in house data*) で、多数の検体が同一変異を保有する場合にはフィルターで除外するようにした。また、2013年末に京都大学が中心に公開した日本人1,208人分のexomeデータ(HGVD)もフィルターに加え、多人数で同一変異が検出されているものは除外した。この後最終的に、DNAから作られるタンパク質のアミノ酸翻訳が止まってしまうnonsense変異、アミノ酸変異を引き起こすmissense変異、アミノ酸翻訳にズレが生じ変異部分以降異なるアミノ酸が翻訳されてしまうframeshift変異に絞ったのち、同一家系で複数検体がある場合には家系内の疾患の有無と変異の出現パターンの遺伝形式が合致する変異のみに絞り込んだ。

以上のような解析パイプラインを構築し、実際に研究班で行っている実験系で次世代シーケンスを行い、その精度を検証することとした。検証には、HapMap Projectで用いられた変異が既知の日本人サンプルを用い、変異同定率・一致率を調査した。全exomeの検証においては、同一検体をOmniExpressExome chipを用いてgenotypingし、一致率から精度を検証した。その後、作製した改良解析パイプラインを用いて、先天性神経疾患検体の次世代シーケンス実験を行い、累計約300名のexome変異探索を行った。

C. 研究結果

カスタム exome および全 exome シーケンスの結果得られたデータの各検体当たりの平均値を図2に示した。リファレンスゲノムへのマッピング率はいずれも99%以上、duplication read (冗長リード)は5~10%程度と良好なデータが得られた。プローブ設計領域の平均読み深度 (depth)

	神経系遺伝子 カスタムexome	All exome
キャプチャープローブ設計領域長	1.6 M bases	51 M bases
シーケンスターゲット領域長 (キャプチャー領域±100bp)	4.4 M bases	88 M bases
シーケンス手法	101 bp paired-end, 24 multiplex / lane	101 bp paired-end, 6 multiplex / lane
シーケンスデータ平均リード長	2.11 G bases	7.88 G bases
ゲノムマッピング率	99.30 %	99.68 %
冗長除去リード率	11.62 %	4.84 %
ユニークにマッピングされpair read間の距離が平均±25Dのリード中、ターゲット領域に含まれるリード長 [率]	1.25 G bases [74.66 %]	6.14 G bases [87.75 %]
プローブ設計領域平均読み深度	1001.9	98.8
シーケンスターゲット領域平均読み深度	811.9	68.8
プローブ設計領域の読み深度10以上 (SNV call可能場所)の割合	99.7 %	98.0 %

図2. Exomeシーケンスの平均データ量結果

はカスタム exome が 1001.9、全 exome が 98.8 と十分な depth が得られていた。実際に SNV を call する際は depth 10 以上の領域のみとしたが、プローブ設計領域のうち depth が 10 以上だった割合は、図2の通りカスタム exome が 99.7%、全 exome が 98.0%であり、ほとんどの設計領域で SNV call 対象領域としてカバーできていた。

この構築した次世代シーケンスデータからの塩基 call システムで得られたシーケンスの精度を確認するため、HapMap プロジェクト

(<http://hapmap.ncbi.nlm.nih.gov>) にて SNP genotyping された日本人 NA18943 検体を我々の実験系でシーケンス call を行い、HapMap プロジェクトとの call の一致度を調査した。まず、カスタム exome に関しては、HapMap で genotype データがある箇所と我々のシーケンスデータがある箇所との間で3,129箇所の重複があった。その3,129箇所の genotype の一致率について調べたところ、HapMap データと我々の次世代シーケンスデータで16箇所だけ不一致であった (一致率 99.49% = 16/3129)。この不一致であった16箇所について、真の genotype は何であることを確認するため、サンガーシーケンス法にて再シーケンスを行ったところ、実際は16箇所全てにおいて我々の genotype call が合っていた (HapMap データが誤っていた。図3)。すなわち HapMap の他の genotype データが全て合っているとしたら、我々の call は極めて100%近く正確であることが確認できた。同様に、全 exome について、HapMap NA18943 検体を OmniExpressExome chip にて genotyping し、全 exome 次世代シーケンスと比較したところ、両方で genotype データが得られている箇所は245,488箇所重複があり、うち162箇所が不一致であった (一致率 99.93% = 162 / 245488)。この

SNV call	position	rsID	Ref	CONGS	HapMap	sanger seq
Homo	chr6: 133,827,354	rs2277083	A	G	A/G	G
Homo	chr9: 123152241	rs2297456	A	C	A/C	C
Hetero	chr8: 144,887,734	rs11783725	G	A/G	A	A/G
Hetero	chr9: 139,328,369	rs10781543	T	C/T	T	C/T
Hetero	chr17: 7,579,472	rs1042522	G	C/G	G	C/G
Ref	chr1: 22,176,959	rs2290499	C	C	T	C
Ref	chr1: 206,574,659	rs1754497	C	C	G	C
Ref	chr3: 12,573,336	rs299641	A	A	G	A
Ref	chr4: 4,861,780	rs10213286	C	C	T	C
Ref	chr6: 151,413,828	rs2295734	C	C	A	C
Ref	chr7: 103,206,013	rs2711839	T	T	A	T
Ref	chr7: 103,244,981	rs362681	A	A	G	A
Ref	chr16: 2,143,849	rs2854581	C	C	G	C
Ref	chr16: 2,162,979	rs2855349	G	G	A	G
Ref	chr19: 41,034,912	rs1461008	G	G	A/G	G
Ref	chr20: 62,038,757	rs3746364	C	C	C/T	C

図3. 我々のデータとHapMapデータで不一致だった箇所のサンガーシーケンス結果。「CONGS」の列が我々の次世代シーケンスでのgenotype call結果を示す。

場合の真の精度はサンガーシーケンス未実施のため上記のようにHapMapがcallエラーの場合も考えられるため少なくとも99.93%以上の精度と考えられると言うに留められる。

実際の先天性疾患の全exome百数十例において、callされたSNVおよびindelの数の検体ごとの平均数を図4に示した。データベース等において既知のSNPを除く前の段階でcallされてくるSNVの数は約7万弱、indelの数は約5,000であった。その後、既知のSNPを除外し、exon領域のnonsense、missense、frameshift変異に絞った上で、in-house検体で多数見られたSNVを除くと、残数は1検体当たりの平均で97 SNVsおよび8 indelとなった。ここから家系ごとに遺伝形式(表現型)に合致するvariantに絞ると、

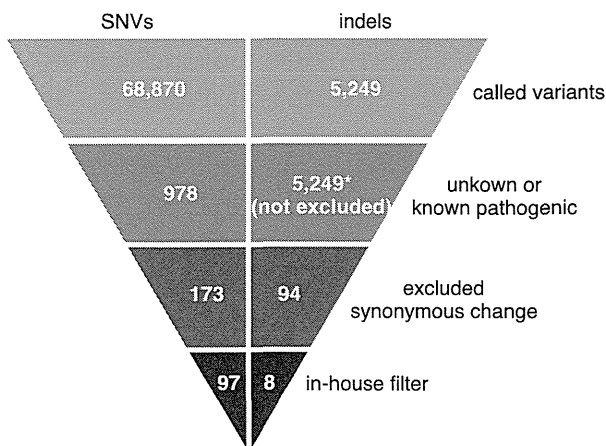


図4. CallされたSNVおよびindelの数と絞り込み段階で残った数の1検体当たり平均数

家系ごとに多少の差はあるが最終的に1~4個以下まで絞れる場合がほとんどであった(疾病ごと

の詳細な進捗状況については研究分担者の各報告を参照)。見つかったほぼ確定的な疾患原因候補遺伝子変異の主な遺伝形式パターンを図5に示した。ほぼ確定的な検体に占める遺伝形式のパターンは、興味深いことに、いわゆる古典的典型的例の両親から同一箇所の変異を受け継

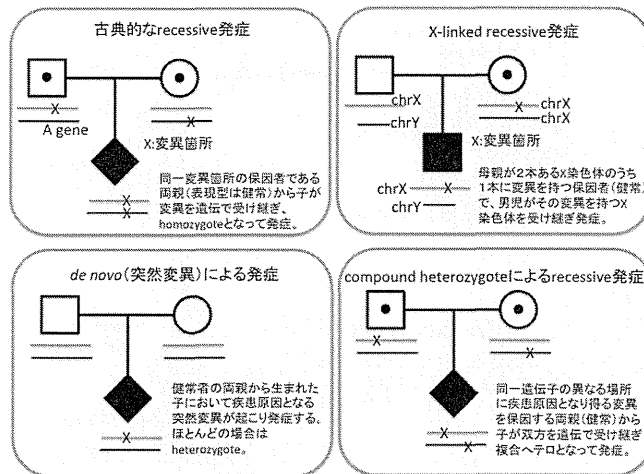


図5. 観察された主な遺伝形式パターン

ぎ、homozygoteで発症していると見られるrecessiveの遺伝形式(図5左上)は全例中わずか1例にしか見られず、突然変異によるde novo mutationのheterozygoteで発症していると考えられる遺伝形式(図5左下)、および同一遺伝子の別の箇所に異なる変異を持つ両親から両変異を受け継ぎ発症していると考えられるcompound heterozygote(図5右下)の2つのパターンが最も多く見られた(その2パターンが各々36.8%)。また、男児患者においてはX染色体連鎖劣性遺伝形式で発症していると考えられる症例が4例見受けられた(割合は男児患者の家族例数により変動する)。さらに興味深いことに、compound heterozygoteと考えられる患者例のうち2例においては、SNVだけではなくindelが関与するcompound heterozygoteが観察された。

以上のように、多数の検体で原因と考えられる変異候補を同定したが、一方で候補が一つも見つからない(絞り込み段階で候補mutationが全て消えてしまう)家系も約15%存在した。様々な要因が考えられるが、次世代シーケンスの全exomeで配列が解析可能であった領域を確認したところ、タンパク質をコードしているexon領域(CDS領域)中、カバーできている領域は93.6%であった。このことから、カバーされていない残りの領域に見つけられていない原因変異があることが考えられたため、我々は今回用いたSureSelectでの全exomeの他に、HaloPlexというゲノムを制限酵素で切断後に特定領域のPCR ampliconを作製し次世代シーケンスするという手法でのカスタムシ

シーケンスを、未カバー領域について追加実施した。CDS の未カバー領域 6.4%のうち HaloPlex によるカスタムプローブを設計できた領域は 67.7%であり、それを用いて実際にシーケンスしたところ variant call に十分な読み深度 (depth 10 以上) が得られた領域は設計領域中 80.8%であった (図 6)。

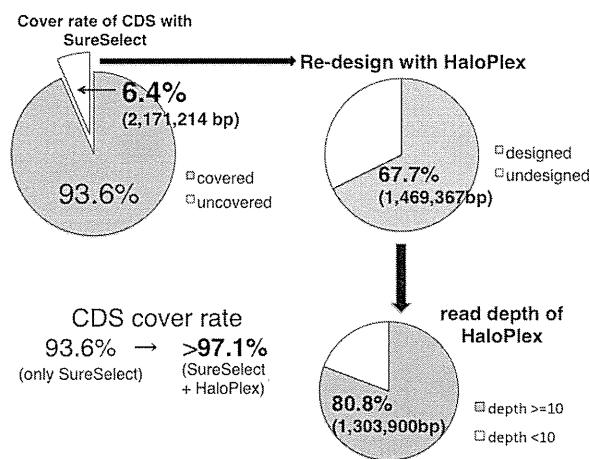


図 6 .追加カスタムHaloPlex exomeによるCDS領域のカバー率

このカスタムexomeにより、CDS領域中のカバー率は合計97.1%以上になった。現在実際の先天性疾患検体にて追加exomeを実施中であるが、これまで候補が見つからなかった検体においても新規の原因候補が見つかってきており、解析を継続している。

D. 考察

我々は先天性疾患の exome 解析のための高精度な variant call 解析システムを開発した。カスタム exome で検証した場合は 3,129 カ所全てが正しく call され偽陽性・偽陰性共に極めて 0 に近いことが推測され、また全 exome の場合でも検証した 245,488 カ所について少なくとも 99.93% 以上は正しい call が得られていることが推測された。これは SNP chip データと不一致だった 0.07% についてサンガーシーケンスによる validation を行なうことによって、おそらくより 100% 近く正しい call が得られていることが証明されたと考えられる。カスタム exome の validation の際に HapMap データと不一致だった 16 カ所全てが我々の次世代シーケンスデータの方が真実の call であったことから、実際には SNP chip の精度よりも我々の次世代シーケンスデータ解析パイプラインによる call の方がよりエラー率が低いと考えられた。

実際の先天性疾患家系の exome 解析により多数の家系での原因遺伝子変異と考えられる変異を同定することができたが、当初想定したよりも compound heterozygote による発症と考えられる例が多数見つかったことが興味深かった。特に

compound heterozygote と考えられる患者例のうち 2 例においては、SNV だけではなく indel が関与する compound heterozygote が観察されたことは、今後の exome 解析において見逃さないように着目すべき点が示唆された例であると考えられる。一般的に SNV に比べると indel は call が難しく高精度な call システムが必要不可欠であるが、我々の例ではサンガーシーケンスにより疾患関連候補 indel は今現在までのところ全て確認されているので、その点でも高精度なシステムができていると考えられる。

また、候補遺伝子変異が一つも残らない家系が我々の系でも 15% 存在したことは今後の課題である。これは様々な要因が考えられ、例えば遺伝子配列だけでは解明できないメチル化等のエピジェネティックな要因による要因、妊娠中の特定の環境要因の曝露による要因、等が挙げられるが、exome 解析で見逃している領域に原因変異が存在することがまずは有力と考えられる。その穴を埋めるために、我々は exome での見逃し領域にフォーカスしたカスタム HaloPlex システムを構築し、全 exome と統合して exon の CDS 領域の 97.1% 以上をカバーできるようになった。現在解析中であるが、この追加カスタム exome で新規の疾患原因候補がいくつか見つかってきており、全 exome でもカバーできていなかったさらなる追加の疾患原因変異が同定できることが期待できる。少なくとも遺伝配列要因による発症原因はこれによりさらに補完できるものと考えている。

今回解析した先天性疾患家系のうち、いくつかについては特定の臨床所見は得られているものの確定診断がついていない家系も含まれていたが、exome 解析により既知の疾患原因変異と同一の変異を見つげられたことで確定診断に至った例が数例存在した。このことは、exome 解析が臨床診断にも非常に有用である例を示している。exome 解析は未知 (新規) の疾患原因遺伝子変異を探索するという基礎科学的な側面と、実際の臨床での確定診断のためのスクリーニングに有用であるという側面のいずれにも有用であろう。神経系先天性疾患の場合、当初の神経系遺伝子の 284 遺伝子のカスタムでは候補が見つからず、全 exome に回した検体も複数あったが、全 exome で見つかった変異をさらにカスタムの作製にフィードバックさせ、より神経系原因遺伝子変異を網羅したカスタム exome を作製できれば、臨床でのスクリーニングにもさらに有用となるシステムが作製できるものと考えられる。これは他の各種疾患も同様で、今後の更なる知見の蓄積で疾患ごとのより効率的で検出力の高いカスタム exome スクリーニングシステムの作製が可能となると予想され、これはそう遠くないうち、不完全ながらも現在からでも作製していくことで臨床の現場に

フィードバックされるようになることが大いに期待される。

E. 結論

我々は先天性疾患コンソーシアムを形成し、原因遺伝子変異未確定の患者およびその家族について、次世代シーケンサーを用いたexome解析を行なうことで、疾患原因となっている遺伝子変異候補を同定する解析システムを完成させた。そのcall精度は少なくとも99.93%以上と推測された。実際にこの解析システムを用いて約300例のexome解析を実施し、数十家系で疾患原因変異と強く示唆される確定的な変異の同定に至った。見つけた疾患原因変異候補の中にはcompound heterozygoteでの発症が疑われる例が、我々の検体中では36.8%と高頻度で散見され、SNVとindelによるcompound heterozygoteも見られ、exome解析の一つの着目点が示された。新規の原因遺伝子候補を複数発見できたという基礎科学的な側面の進展と共に、疾患名未確定の検体において既知の原因変異の同定から確定診断に至るといった臨床的な有用性例も示すことができた。全exome解析でカバーされていない領域をさらに補完するカスタムexomeを開発し、全exonのCDS領域中97.1%以上をカバーできるシステムを構築した。今後臨床の現場にexome解析がさらに利用されていくことが期待される。

F. 研究発表

1. 論文発表

- 1) Tateno C, Miya F, Wake K, Kataoka M, Ishida Y, Yamasaki C, Kakuni M, Wisse E, Verheyen F, noue K, Sato K, Kudo A, Arii S, Itamoto T, Asahara T, Tsunoda T, Yoshizato K; Morphological and microarray analyses of human hepatocytes from xenogeneic host livers. *Lab. Invest.* **93**, 54-71 (2013).
- 2) Fujimoto A, Totoki Y, Abe T, Boroevich KA, Hosoda F, Nguyen HH, Aoki M, Hosono N, Kubo M, Miya E, Arai Y, Takahashi H, Shirakihara T, Nagasaki M, Shibuya T, Nakano K, Watanabe-Makino K, Tanaka H, Nakamura H, Kusuda J, Ojima H, Shimada K, Okusaka T, Ueno M, Shigekawa Y, Kawakami Y, Arihiro K, Ohdan H, Gotoh K, Ishikawa O, Ariizumi S, Yamamoto M, Yamada T, Chayama K, Kosuge T, Yamaue H, Kamatani N, Miyano S, Nakagama H, Nakamura Y, Tsunoda T, Shibata T, Nakagawa H; Whole-genome sequencing of liver cancers identifies etiological influences on mutation patterns and recurrent mutations in chromatin regulators. *Nat. Genet.* **44**, 760-764 (2012).
- 3) Tsuge M, Fujimoto Y, Hiraga N, Zhang Y, Ohnishi M, Kohno T, Abe H, Miki D, Imamura M, Takahashi S, Ochi H, Hayes CN, Miya E,

Tsunoda T, Chayama K; Hepatitis C virus infection suppresses the interferon response in the liver of the human hepatocyte chimeric mouse. *PLoS One* **6**, e23856 (2011).

- 4) Tateno C, Kataoka M, Utoh R, Tachibana A, Itamoto T, Asahara T, Miya F, Tsunoda T, Yoshizato K; Growth Hormone-dependent Pathogenesis of Human Hepatic Steatosis in a Novel Mouse Model Bearing a Human Hepatocyte-repopulated Liver. *Endocrinology* **152**, 1479-1491 (2011).
 - 5) Kato Y, Zembutsu H, Takata R, Miya F, Tsunoda T, Obara W, Fujioka T, Nakamura Y; Predicting response of bladder cancers to gemcitabine and carboplatin neoadjuvant chemotherapy through genome-wide gene expression profiling. *Exp. Ther. Med.* **1**, 47-56 (2011).
 - 6) Okada Y, Miya F, Koike M, Tomisato S, Tokura T, Ishihara Y, Shimojo D, Kanematsu D, Kanemura Y, Kohda K, Sobue G, Yamanaka S, Yuzaki M, Uchiyama Y, Ikeda E, Tsunoda T, Okano H; Incompletely reprogrammed human induced pluripotent stem cells form glioma-like tumors through genomic instability during neural differentiation. *Cell Stem Cell (submitted 2013)*
 - 7) Hirokawa M, Tajima T, Takahashi A, Ashikawa K, Miya F, Shigemizu D, Ozaki K, Sakata Y, Nakatani D, Suna S, Imai Y, Tanaka T, Tsunoda T, Matsuda K, Kadowaki T, Nakamura Y, Nagai R, Komuro I, Kubo M; A genome-wide association study identifies PLCL2 and AP3D1-DOT1L-SF3A2 as new susceptibility loci for myocardial infarction in Japanese. *Eur. J. Hum. Genet. (Submitted 2014)*
- ### 2. 学会発表
- 1) Miya E, Morizono T, Abe T, Kubo M, Tsunoda T; Massive Genome-wide cis- and trans-eQTL analysis using next generation sequencer in Japanese population. The 36th Annual Meeting of the Molecular Biology Society of Japan, Kobe (2013).
 - 2) 宮 冬樹, 森園 隆、阿部 哲雄、久保 充明、角田 達彦; 日本人における全ゲノム間の関連についての網羅的eQTL解析 (Massive Genome-wide cis- and trans-eQTL analysis in Japanese population), 日本人類遺伝学会 第58回大会, 仙台 (2013).
 - 3) 岡本 奈那, 岡本 伸彦, 川戸 和美, 松田 圭子, 三島 祐子, 山本 悠斗, 宮 冬樹, 角田 達彦, 加藤 光広, 齋藤 伸治, 山崎 麻美, 金村 米博, 小崎 健次郎; 神経疾患を標的にした次世代シーケンサー解析で診断したBaraitser-Winter症候群の1例, 日本人類遺伝学会 第58回大会, 仙台 (2013).
 - 4) Okamoto N, Miya E, Tsunoda T, Kato M,

- Saitoh S, Yamasaki M, Kanemura Y, Kosaki K; Application of targeted next-generation sequencing in the diagnosis of pediatric neurological disorders. American Society of Human Genetics, Boston (2013).
- 5) Okada Y, Miya E, Koike M, Tomisato S, Tokura T, Ishihara Y, Shimojo D, Hattori C, Kanematsu D, Kanemura Y, Kohda K, Sobue G, Yamanaka S, Yuzaki M, Uchiyama Y, Ikeda E, Tsunoda T, Okano H; Incompletely reprogrammed human iPSCs form glioma-like tumors through genomic instability during neural differentiation. The 36th Annual Meeting of the Japan Neuroscience Society, Kyoto (2013).
 - 6) 服部 文子, 根岸 豊, 宮 冬樹, 安藤 直樹, 伊藤 哲哉, 角田 達彦, 金村 米博, 小崎 健次郎, 齋藤 伸治; AKT3遺伝子変異による巨脳症の一例 (A case with megalencephaly caused by an ATK3 mutation), 第55回日本小児神経学会学術集会, 大分 (2013).
 - 7) Miya E, Morizono T, Abe T, Kubo M, Tsunoda T; Massive genome-wide eQTL analysis in Japanese population. The 35th Annual Meeting of the Molecular Biology Society of Japan, Fukuoka (2012).
 - 8) Okada Y, Miya E, Koike M, Tomisato S, Kanematsu D, Kanemura Y, Kohda K, Yamanaka S, Yuzaki M, Uchiyama Y, Ikeda E, Okano H; Incompletely reprogrammed human iPSCs form glioma-like tumors through genomic instability during neural differentiation. The 35th Annual Meeting of the Molecular Biology Society of Japan, Fukuoka (2012).
 - 9) 宮 冬樹, 阿部 哲雄, 森園 隆, 秋山 真太郎, 久保 充明, 角田 達彦; 次世代シーケンサーを用いた日本人における全ゲノム網羅的 eQTL 解析, 日本人類遺伝学会 第57回大会, 東京 (2012).
 - 10) 鈴木 直大, 務台 英樹, 鳥居 千春, 清水 厚志, 宮 冬樹, 難波 一徳, 工藤 純, 小崎 健次郎, 松永 達雄; カスタムカーゲットリシーケンスによる難聴関連遺伝子の変異探索, 日本人類遺伝学会 第57回大会, 東京 (2012).
 - 11) 根岸 豊, 服部 文子, 安藤 直樹, 宮 冬樹, 角田 達彦, 北井 征宏, 荒井 洋, 岡本 伸彦, 加藤 光広, 山崎 麻美, 金村 米博, 小崎 健次郎, 齋藤 伸治; 網羅的遺伝子解析により難治性てんかん患児に同定された複合ヘテロSCN2A遺伝子変異, 日本人類遺伝学会 第57回大会, 東京 (2012).
 - 12) Okada Y, Miya E, Kanemura Y, Koike M, Kohda K, Yuzaki M, Uchiyam Y, Tsunoda T, Yamanaka S, Okano H. Evaluation of human iPSC cells by neural differentiation and tumorigenicity. 8th IBRO World congress of Neuroscience, Florence (2011).
 - 13) Ishii S, Okada Y, Miya E, Tsunoda T, Shimazaki T, Okano H. Efficient generation and developmental analysis of basal forebrain cholinergic neurons from mouse embryonic stem cells. 8th IBRO World congress of Neuroscience, Florence (2011).
 - 14) Ishii S, Okada Y, Miya E, Tsunoda T, Shimazaki T, Okano H. Efficient generation and developmental analysis of basal forebrain cholinergic neurons from mouse embryonic stem cells. The 34th Annual Meeting of the Japan Neuroscience Society, Yokohama (*Neurosci. Res.* **71**, e332) (2011).
 - 15) Okada Y, Miya E, Kanemura Y, Sunabori T, Koike M, Kohda K, Yuzaki M, Uchiyama Y, Tsunoda T, Yamanaka S, Okano H, Evaluation of human iPSC cells by neural differentiation and tumorigenicity. The 34th Annual Meeting of the Japan Neuroscience Society, Yokohama (*Neurosci. Res.* **71**, e329-e330) (2011).
 - 16) Ishii S, Okada Y, Miya E, Tsunoda T, Shimazaki T, Okano H.: Efficient generation and developmental analysis of basal forebrain cholinergic neurons from mouse embryonic stem cells. Society for Neuroscience 41th Annual Meeting, Washington DC (2011)
 - 17) 加藤 陽一郎, 前佛 均, 高田 亮, 角田 達彦, 宮 冬樹, 小原 航, 中村 祐輔, 藤岡 知昭; 浸潤性膀胱癌に対するカルボプラチン/ジェムシタビン (CaG) の感受性予測システムの構築, 第99回日本泌尿器科学会総会, 名古屋 (2011).
- G. 知的財産権の出願・登録状況 (予定を含む。)
1. 特許取得
 - 1) METHOD FOR SELECTING IPS CELL CLONE, AND METHOD FOR SELECTING GENE USED IN METHOD FOR SELECTING SAME (iPS細胞クローンの選択方法、及びその選択方法に用いる遺伝子の選択方法), 特許出願, 世界知的所有権機関(WIPO) 国際公開番号: WO 2012/115270, 日本出願番号: 特願 2011-040979, アメリカ出願番号: 201161466298P, 出願人: Yohei Okada (岡田 洋平), Hideyuki Okano (岡野 栄之), Tatsuhiko Tsunoda (角田 達彦), Fuyuki Miya (宮 冬樹). (2012)
 - 2) ヒト肝細胞を有するマウスの処置方法. 日本国特許取得, PATENT NUMBER 特許第5025173号 (出願番号 2006-181621, 公開番号: 2007-037542, 登録日: 2012.6.29), 国際出願(PCT) 済 出願番号: PCT/JP2006/313077, 出願人: 向谷 知世, 吉里 勝利, 舛本 法生, 片岡 美穂, 角田 達彦, 宮 冬樹. (2012)

2. 実用新案登録
なし

3. その他
なし