

臨床ゲノム情報を活用した 診断支援システムの要素技術の検討

高野 良治^{*1}, 宮本 青^{*1}, 北島 正人^{*2}, 赤堀 光希^{*1},
石浦 浩之^{*3}, 辻 省次^{*4,5}, 大江 和彦^{*4}

^{*1} 富士通株式会社, ^{*2} 株式会社富士通九州システムズ,

^{*3} 東京大学医学部附属病院 神経内科, ^{*4} 東京大学大学院 医学系研究科,

^{*5} 国際医療福祉大学 医学研究科

Rapid prototyping of genomic diagnostic platform for clinical decision support

Takano Ryohji ^{*1}, Miyamoto Sei ^{*1}, Kitajima Masato ^{*2}, Akahori Kouki ^{*1},

Ishiura Hiroyuki ^{*3}, Tsuji Shoji ^{*4,5}, Ohe Kazuhiko ^{*4}

^{*1} Fujitsu Limited, ^{*2} Fujitsu Kyushu Systems Limited, ^{*3} The University of Tokyo Hospital,

^{*4} Graduate School of Medicine and Faculty of Medicine, The University of Tokyo,

^{*5} Institute of Medical Genomics, International University of Health and Welfare.

次世代シーケンサーの普及に伴い臨床現場における患者のゲノム情報の活用が始まっている。一方、診断に資する情報の抽出・管理や機微な情報の取り扱いなど検討すべき課題も顕在化してきている。我々は、将来的な全ゲノム情報を用いた医療の実現も見据え、診断支援の上で重要になるとと思われる 4 つのシステム課題(①大規模なデータの管理、②効率的な病的変異の特定、③データ品質・信頼性の確保、④個人情報保護への配慮)に関して、プロトタイピングによる概念実証を通じ、対策方法やその実現性、効果などを検証した。その結果、膨大なゲノム情報を管理するためのデータベース(DB)の実装方式、効率的に病的変異を特定するためのユーザインターフェース、ゲノム情報の品質・信頼性を確保するための情報管理の仕組み、個人情報に配慮したゲノムレポートの管理・参照方法に関して、臨床現場での実装に向けた有効性を確認し、その一部について先進医療の中で運用を開始した。

● キーワード: ゲノム医療, 全ゲノムデータ処理, 診断支援システム, 病的変異の特定, 電子カルテ連携

1. はじめに

次世代シーケンサーの普及に伴い患者のゲノム情報に基づく個別化医療が臨床の現場で実践され始めている(Fig.1)。

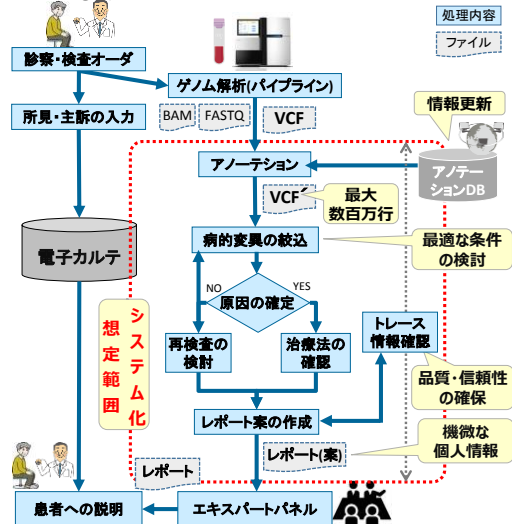


Fig.1 ゲノム情報を活用した医療のフロー

近い将来には全ゲノム情報を用いた検査や診断も想定され、以下のような課題への対応が急務になる。

1) 大規模なデータの管理

1 患者あたり最大数百万箇所に基づくゲノムバリエーション情報と、バリエーションに対する最大数百項目のアノテーション情報を一元的に管理する基盤が必要となる。

2) 効率的な病的変異の特定

最大数百万に及ぶ膨大なバリエーションから、臨床所見や生物・医学的アノテーション情報を用いて疾患の原因や治療標的となるバリエーション(病的変異)を精確かつ効率的に絞り込む必要がある。

3) データ品質・信頼性の確保

診断結果は、解析パイプラインやアノテーションのバージョンなど、データ処理の条件に依存して大きく変動する可能性があるため、ゲノムデータの品質や解析の信頼性を担保するための仕組みが必要となる。

4) 個人情報保護への配慮

ゲノム情報それ自体が機微な個人情報に該当するため、個人情報保護に配慮した仕組みが必要となる。

我々は、Fig.1 の点線内の業務(ゲノム情報を用いた診断支援)を、システム化により効率化することを目指し、プロトタイピングによる概念実証を通じて上記 4 点の課題解決に向けたシステムの実現性や有効性を検証した。

2. 方法

1) 「大規模なデータの管理」に関する検証

ゲノムバリエーションを一意に特定する「ゲノムキー」を設

計・導入することで患者間に重複するバリエントを排除し、従来の患者毎のデータ管理方法に対するデータ量の削減効果を日本人のバリエントに関する参照データ[1]を用いて検証した。次にゲノムデータの管理に最適な DBMS 選定のため、1000 名の仮想患者のゲノムデータ(50 万変異/患者)を作成し、項目拡張性やスケールアウトの観点から性能の比較検証を行った。

2) 「効率的な病的変異の特定」に関する検証

ClinVar や HGMD などアノテーション DB によって付与される情報(アレル頻度、データ品質指標、遺伝子名、病原性指標等)を利用して効果的に病的変異を絞込む条件や項目を検討した。また、そこから病的変異を特定しゲノムレポートを作成する工程までを含めた一連の処理を効率化するために Web システムのプロトタイプを構築し最適な機能や GUI を検証した。

3) 「データ品質・信頼性の確保」に関する検証

ゲノムデータのトレーサビリティや解析の再現性、正確性を担保するために必要な項目やパラメータを、解析パイプラインを構成するツールや DB、ファイル仕様から精査し、モック開発を通じて医療従事者に提示するコンテンツやユースケースを検証した。

4) 「個人情報保護への配慮」に関する検証

機微な情報が記載されるゲノムレポート(PDF)の安全管理・参照を実現するために、電子カルテから匿名化された SS-MIX2 拡張ストレージへ Web ブラウザを介してアクセスする仕組みのプロトタイプを構築し、システム構成やデータロジスティクスを検証を行った。

3. 結果

1) 「大規模なデータの管理」に関する検証結果

1000 名の患者のゲノム変異を管理する場合、ゲノムキーの適用により患者(変異ファイル)単位で管理する場合と比較管理データを 1%以下に削減できることを示した。また、大規模ゲノムデータの一元的な管理においては分散型 DB が適しているとの結果を得た。

2) 「効率的な病的変異の特定」に関する検証結果

アノテーションデータの精査の結果、9 つの大項目から階層的に絞り込み条件を設定可能な病的変異の検索インターフェースを実装した(Fig.2)。

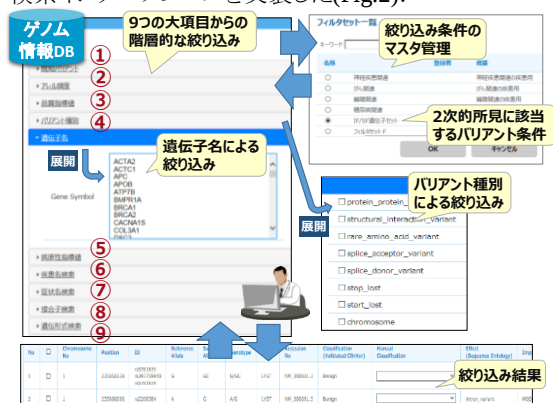


Fig.2 病的変異の絞り込みインターフェース

複数のアノテーションソースから得られる同種の検索項目は同一階層に配置して検索条件の検討を容易にし、設定した検索条件をマスター管理可能な構成とすることで、疾患や診療科毎の絞り込み条件の再利用や業務効率化に寄与することを確認した。さらに二次的所見とすべきバリエントの絞り込み条件をマスター登録し

ておくことで、患者に返すべき情報とそうでない情報を区別する仕組みとしても活用し得ることを示した。

3) 「データ品質・信頼性の確保」に関する検証結果

ゲノムレポートを作成した時点における諸条件(DB バージョン、解析パラメーター等)をデータ品質レポートとして生成し、医療従事者がゲノムレポートを参照するタイミングで、データ品質に関する確認を選択できるようにすることで、検体だけでなくデータの精度管理にも貢献し得ることを確認した。また、アノテーション DB の更新に伴う重要な変更情報をモニタリングし、医療従事者へアラート通知する機能をもたせることで、頻繁に更新されるアノテーション情報の変更を確認する負担が軽減され、最新の情報に基づく診断結果が患者に返される確度が高まることが期待された。

4) 「個人情報保護への配慮」に関する検証結果

対応表を用いて実患者 ID と匿名 ID を変換・管理し、レポートを検体や検査区分(先進医療、臨床研究等)毎に識別する ID を導入した上で、医療従事者の役割に応じたアクセス権限を設定することで、必要な医療従事者のみが電子カルテからゲノムレポートを参照できる仕組みを実現した。またその一部を先進医療(がんゲノム医療)において実運用するに至った(Fig.3)。さらに、SS-MIX2 拡張ストレージや汎用的な Web システムの仕様としたことで、今後のゲノム医療(保険診療)の波及や均てん化の一助となることも期待された。

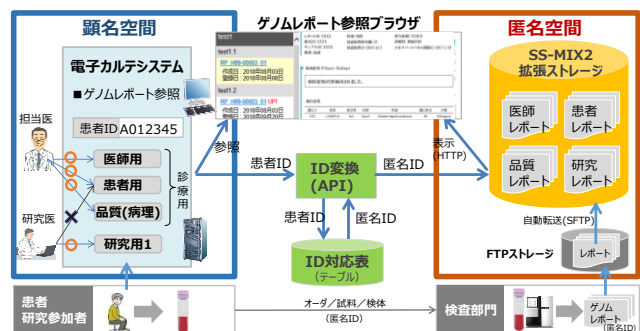


Fig.3 ゲノムレポート参照システム

4. 考察

病的変異特定のプロセスは、最終的に人による解釈やエキスパートパネルでの判断が必要となるため、この点がゲノム検査プロセス全体のボトルネックになる可能性が高い。解釈の均質化や自動化に向けてさらなるシステム化検討が必要と考えられる。

5. 結語

本研究では全ゲノム情報の臨床活用に必要な技術の検討・検証を行い、一部を実運用に適用した。今後は臨床情報の積極的な活用による病的変異の推定や、病情報システムの中での連携・利活用の仕組みを最適化していく必要がある。

参考文献

- [1] Masao Nagasaki, Jun Yasuda, Fumiki Katsuoka, et.al: Rare variant discovery by deep whole-genome sequencing of 1,070 Japanese individuals, Nature Communications, 6, 1-13, 2015.