

科学研究費助成事業 研究成果報告書

平成 28 年 6 月 17 日現在

機関番号：17301

研究種目：基盤研究(C) (一般)

研究期間：2013～2015

課題番号：25430183

研究課題名(和文) ヒト全ゲノムシーケンスのための統合アノテーションワークフローの構築

研究課題名(英文) Development of Integrated Annotation Workflows for Human Whole Genome Sequencing

研究代表者

三嶋 博之 (MISHIMA, Hiroyuki)

長崎大学・原爆後障害医療研究所・助教

研究者番号：10513319

交付決定額(研究期間全体)：(直接経費) 4,100,000円

研究成果の概要(和文)：本研究は大規模シーケンサーによる疾患責任変異の探索のためのワークフローの構築を目的とする。このために本研究では公共サービスであるTogoWSの拡張によるUCSC Genome Databaseのサポート、各種オープンソースツール公開、Bio-Virtuosoの公開を行った。また、12名の全ゲノムシーケンシングデータにワークフローを応用した。

研究成果の概要(英文)：This study aims to build workflows for finding disease causing mutations using whole genome sequencing data. For the goal, the primary researcher extended TogoWS to support the UCSC genome database, several open-source tools, the Bio-Virtuoso Docker containers. Built workflows were applied for whole genome dataset of 12 individuals.

研究分野：人類遺伝学

キーワード：ヒトゲノム バイオインフォマティクス 次世代シーケンサー ワークフロー

1. 研究開始当初の背景

大規模シーケンサーによる疾患責任変異の探索は、急速なコストの低下に伴い、全エクソーム解析から全ゲノム解析へと移行しはじめていた。しかし、全ゲノムデータに對して最新のデータベース (DB) 情報を統合して疾患責任候補変異を絞り込むためのワークフローで評価が定まったものは存在しなかった。

2. 研究の目的

本研究は複数の腫瘍データベース情報に容易にアクセスするための基盤の拡張整備を行った上で、新規にアノテーションワークフローを構築し、さらに実際の疾患全ゲノムデータを用いてその有用性を実証することを目的とする。

3. 研究の方法

複数の公共データベースにアクセスするための統一したインターフェース (REST インターフェース) を提供する TogoWS を代表的な大規模データベースのひとつである UCSC ゲノム DB を対応させた。問題点は、UCSC ゲノム DB が 4 万を越える多数のテーブルから構成されることであったが、これまで代表研究者が開発を続けている Ruby UCSC API を内部で用いることでこれを解決した。

多数のデータベースを組み合わせる場合、特に知識情報のようなオントロジー (階層化された語彙集) に基づく情報を効果的に組み合わせるためには、セマンティックウェブ技術の応用し、情報を RDF 形式に統一的に表現し、SPARQL 問い合わせ言語を用いて検索することが効率的である。複数のデータベースからのデータ取得・RDF 化・トリプルストアへの格納を一元化するための仕組みとして Bio-Virtuoso を開発している。

全ゲノムシーケンシングは 4 トリオ家系 (12 名) について全エクソームおよび全ゲノムシーケンシングを完了し、GATK・XHMM・CNVnator などのツールと新規作成ツールとを組み合わせたワークフローを作成した。

4. 研究成果

機能を拡張した TogoWS はすでに公開されており、国内外の研究者に広く用いられている。Bio-Virtuoso は開発中であるが、Git Hub においてオープンソースソフトウェアとして公開されている。ワークフローを構成するソフトウェアは Git Hub において公開されている。特に京都大学 1200 人エクソームアレル頻度情報 HGVD のフォーマット変換ツールは疾患ゲノム解析に広く用いられている。

作成したツールをいくつかのプロジェクトにおいて実際に用いた結果を論文として発表できた。全ゲノムシーケンシングの解析結果については、現在論文準備中である。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

(雑誌論文) (計 13 件)

1. Uchiyama Y, Nakashima M, Watanabe S, Miyajima M, Taguri M, Miyatake S, Miyake N, Saitsu H, Mishima H, Kinoshita A, Arai H, Yoshiura K, Matsumoto N: Ultra-sensitive droplet digital PCR for detecting a low-prevalence somatic GNAQ mutation in Sturge-Weber syndrome. *Scientific Reports* 2016, 6:22985. DOI: 10.1038/srep22985, 査読有
2. Yamamoto T, Mishima H, Mizukami H, Fukahori Y, Umehara T, Murase T, Kobayashi M, Mori S, Nagai T, Fukunaga T, Yamaguchi S, Yoshiura K, Ikematsu K: Metabolic autopsy with next generation sequencing in sudden unexpected death in infancy: Postmortem diagnosis of fatty acid oxidation disorders. *Molecular Genetics and Metabolism Reports* 2015, 5:26-32. DOI: 10.1016/j.ymgmr.2015.09.005, 査読有
3. Tomoshige K, Matsumoto K, Tsuchiya T, Oikawa M, Miyazaki T, Yamasaki N, Mishima H, Kinoshita A, Kubo T, Fukushima K, Yoshiura K, Nagayasu T: Germline mutations causing familial lung cancer. *J Hum Genet* 2015, 60:597-603. DOI: 10.1038/jhg.2015.75, 査読有
4. Tamura S, Higuchi K, Tamaki M, Inoue C, Awazawa R, Mitsuki N, Nakazawa Y, Mishima H, Takahashi K, Kondo O, Imai K, Morio T, Ohara O, Ogi T, Furukawa F, Inoue M, Yoshiura K, Kanazawa N: Novel compound heterozygous DNA ligase IV mutations in an adolescent with a slowly-progressing radiosensitive-severe combined immunodeficiency. *Clinical Immunology*, 2015, 160(2):225-260. DOI: 10.1016/j.clim.2015.07.004, 査読有
5. Miura K, Higashijima A, Mishima H, Miura S, Kitajima M, Kaneuchi M, Yoshiura K, Masuzaki H: Pregnancy-associated microRNAs in plasma as potential molecular markers of ectopic pregnancy. *Fertility and Sterility* 2015, 103:1202-1208.e1. DOI: 10.1016/j.fertnstert.2015.01.041, 査読有
6. Miura K, Morisaki S, Abe S, Higashijima A, Hasegawa Y, Miura S, Tateishi S, Mishima H, Yoshiura K, Masuzaki H: Circulating levels of maternal plasma cell-free pregnancy-associated placenta-specific microRNAs are associated with placental weight. *Placenta* 2014, 35:848-851. DOI: 10.1016/j.placenta.2014.06.002, 査読有

7. Miura K, Hasegawa Y, Abe S, Higashijima A, Miura S, Mishima H, Kinoshita A, Kaneuchi M, Yoshiura K, Masuzaki H: Clinical applications of analysis of plasma circulating complete hydatidiform mole pregnancy-associated miRNAs in gestational trophoblastic neoplasia: A preliminary investigation. *Placenta* 2014, 35:787-789. DOI: 10.1016/j.placenta.2014.06.004, 査読有
8. Miura K, Mishima H, Kinoshita A, Hayashida C, Abe S, Tokunaga K, Masuzaki H, Yoshiura K: Genome-wide association study of HPV-associated cervical cancer in Japanese women. *J. Med. Virol.* 2014, 86:1153-1158. DOI: 10.1002/jmv.23943, 査読有
9. Abe S, Miura K, Kinoshita A, Mishima H, Miura S, Yamasaki K, Hasegawa Y, Higashijima A, Jo O, Yoshida A, Kaneuchi M, Yoshiura K, Masuzaki H: Single human papillomavirus 16 or 52 infection and later cytological findings in Japanese women with NILM or ASC-US. *J Hum Genet* 2014, 59:251-255. DOI: 10.1038/jhg.2014.9, 査読有
10. Tsukamoto O, Miura K, Mishima H, Abe S, Kaneuchi M, Higashijima A, Miura S, Kinoshita A, Yoshiura K, Masuzaki H: Identification of endometrioid endometrial carcinoma-associated microRNAs in tissue and plasma. *Gynecologic Oncology* 132:715-721, 2014. DOI: 10.1016/j.ygyno.2014.01.029, 査読有
11. Miura K, Higashijima A, Miura S, Mishima H, Yamasaki K, Abe S, Hasegawa Y, Kaneuchi M, Yoshida A, Kinoshita A, Yoshiura K, Masuzaki H: Predominantly placenta-expressed mRNAs in maternal plasma as predictive markers for twin-twin transfusion syndrome. *Prenatal Diagnosis* 34:345-349, 2014. DOI: 10.1002/pd.4307, 査読有
12. Abe S, Miura K, Kinoshita A, Mishima H, Miura S, Yamasaki K, Hasegawa Y, Higashijima A, Jo O, Sasaki K, Yoshida A, Yoshiura K, Masuzaki H: Copy number variation of the antimicrobial-gene, defensin beta 4, is associated with susceptibility to cervical cancer. *J Hum Genet* 58:250-253, 2013. DOI: 10.1038/jhg.2013.7, 査読有
13. Higashijima A, Miura K, Mishima H, Kinoshita A, Jo O, Abe S, Hasegawa Y, Miura S, Yamasaki K, Yoshida A, Yoshiura K, Masuzaki H: Characterization of placenta-specific microRNAs in fetal growth restriction pregnancy. *Prenatal Diagnosis* 2013, 33:214-222. DOI: 10.1002/pd.4045, 査読有
- 〔学会発表〕(計7件)
1. 【国内ポスター】三嶋博之, 吉浦孝一郎: Exomiserを用いたエクソーム解析ワークフローと支援ツールの開発. 日本人類遺伝学会第60回大会, 2015年10月14-17日, 東京都京王プラザホテル
 2. 【国内ポスター】三嶋博之, 吉浦孝一郎: (1-2) Exomiserを用いた知識ベース疾患原因変異探索のためのワークフロー. NGS現場の会第四回研究会, 2015年7月1-3日, つくば市つくば国際会議場
 3. 【国内口演】三嶋博之, 菊入崇, 三古谷忠, 吉浦孝一郎: (101-3) 多発性歯牙腫合併症例を含むSATB2遺伝子変異症候群の新規変異の同定. 日本人類遺伝学会第59回大会, 2014年11月20日, 東京都江戸川区タワーホール深堀
 4. 【国内ポスター】Hiroyuki Mishima, Koh-ichiro Yoshiura: Handmaid Building of a High-Performance Computing Cluster and a Storage System for Bioinformatics. 生命医薬情報学連合大会, 2014年10月2-4日, 仙台市仙台国際センター.
 5. 【国内口演】三嶋博之, 西澤達也, 吉浦孝一郎, 片山俊明: (044) TogoWS RESTサービスによるUCSCゲノムデータベースの利用. 日本人類遺伝学会第58回大会, 2013年11月21日, 仙台市江陽グランドホテル
 6. 【国内口演】三嶋博之: Gene Hunting: できること・できないこと(フォーカストセッション・オミックス解析における実務者意見交換会). 生命医薬情報学連合大会, 2013年10月29日, 東京都江戸川区タワーホール深堀
 7. 【国内ポスター】三嶋博之, 西澤達也, 片山俊明: (2-47-B) TogoWS RESTサービスによるUCSCゲノムデータベースの利用. NGS現場の会第三回研究会, 2013年9月4-5日, 神戸市神戸国際会議場
- 〔図書〕(計4件)
1. 三嶋博之・清水厚志: コマンドラインの使い方, 学研メディカル秀潤社, 2015年, 細胞工学別冊・次世代シーケンサーDRY解析教本, 23-42
 2. 三嶋博之: UCSCゲノムブラウザ, 羊土社, 2014年, 実験医学増刊・今日から使えるデータベース・ウェブツール達人になるための実線ガイド100, 実験医学 32(20): 3223-3225
 3. 三嶋博之: さまざまなヒトバリエーショ

- ンデータベース, 羊土社, 2014年, 実験
医学増刊・今日から使えるデータベー
ス・ウェブツール達人になるための実践
ガイド 100, 実験医学 32(20): 3341-3343
4. 三嶋博之: 全エクソーム解析における情
報処理, 医歯薬出版, 2013年, 医学のあ
ゆみ 245(5): 345-351

〔産業財産権〕
出願状況(計0件)

名称:
発明者:
権利者:
種類:
番号:
出願年月日:
国内外の別:

取得状況(計0件)

名称:
発明者:
権利者:
種類:
番号:
取得年月日:
国内外の別:

〔その他〕
ホームページ等

Togo WS

<http://www.togows.org/>

Hgvd2annovar

<https://github.com/misshie/hgvd2annovar>

Bio - Virtuoso

<https://github.com/misshie/bio-virtuoso>

アウトリーチ活動

1. 【パネリスト・講師】三嶋博之: ゲノム
データの利活用で暮らしがよくなるた
めには. MIT メディアラボ×朝日新聞
「未来メディア塾 2015」, 2015年10月
18日, 東京都江東区日本科学未来館.

報道関連

1. 伊藤穰一, 加藤崇, 小沼大地, 三嶋博之,
御手洗瑞子: MIT メディアラボ×朝日新
聞 未来メディア塾 2015, 2015年
10月29日, 朝日新聞全国版 27面,
[http://www.asahi.com/articles/DA3S1
2040065.html](http://www.asahi.com/articles/DA3S12040065.html)

6. 研究組織

(1) 研究代表者

三嶋 博之 (MISHIMA, Hiroyuki)

長崎大学・原爆後障害医療研究所・助教

研究者番号: 10513319

(3) 連携研究者

吉浦 孝一郎 (YOSHIURA, Koh-ihiro)

長崎大学・原爆後障害医療研究所・教授

研究者番号: 00304931

片山 俊明 (KATAYAMA, Toshiaki)

情報・システム研究機構 DBCLS・助教

研究者番号: 60396869