

**科学研究費助成事業 研究成果報告書**

平成 30 年 5 月 16 日現在

機関番号：13901

研究種目：挑戦的萌芽研究

研究期間：2015～2017

課題番号：15K12139

研究課題名(和文)機能性ncRNA探索のための転写調節ネットワークアトラス

研究課題名(英文)Transcription regulatory network atlas for functional ncRNA search

研究代表者

島村 徹平(Shimamura, Teppei)

名古屋大学・医学系研究科・特任准教授

研究者番号：00623943

交付決定額(研究期間全体)：(直接経費) 2,800,000円

研究成果の概要(和文)：本研究課題では、転写制御因子とそのターゲット遺伝子群の制御関係が調整因子(モジュレーター)依存的に変化するダイナミクスを記述した確率モデルを観測データから推定し、さまざまな転写制御因子の転写活性を制御するncRNAを網羅的に同定する。これにより、機能性ncRNAを効率的に探索するための転写調節ネットワークアトラスを構築した。シミュレーションデータおよび実データに本開発手法を適用した結果、以下のような成果を得た。(A)がん遺伝子 MYC を制御するノンコーディング RNA の発見、(B)モジュレーター因子の網羅的探索のための解析パイプラインの開発、および転写調節ネットワークアトラスの構築。

研究成果の概要(英文)：In this research project, the purpose is to derive a probabilistic model describing the dynamics in which the regulatory relationship between the transcriptional regulatory factor and its target gene group is changed depending on the modulator from observed expression data and to identify functional ncRNAs controlling the transcriptional activity of various transcriptional regulators. As a result of applying the proposed method to simulation data and actual data, the following results were obtained. (A) Discovery of noncoding RNA that controls the oncogene MYC, and (B) development of a large-scale analysis pipeline for the comprehensive search of modulator factors, and construction of a transcriptional regulatory network atlas.

研究分野：バイオインフォマティクス

キーワード：バイオインフォマティクス システム生物学 遺伝子ネットワーク 発現制御

### 1. 研究開始当初の背景

ヒトゲノムのうち、たんぱく質をコードする領域はわずか2%弱にすぎない。残り98%の「ジャンク」と考えられていたゲノムの広大なノンコーディング領域から、膨大な数の ncRNA が転写され、さまざまな生命現象において重要な役割を果たしていることが徐々に明らかになっている。しかしその一方で、膨大な数にのぼるノンコーディング RNA (ncRNA) の作用機構の全貌は未だ見えない。このような背景から、ncRNA の大海原で研究者が遭難しないよう「予測地図」を提供し、機能性 ncRNA 研究に劇的な変化をもたらす解析プラットフォームの開発が必要である。

### 2. 研究の目的

本研究課題では、転写制御因子とそのターゲット遺伝子群の制御関係が調整因子(モジュレーター)依存的に変化するダイナミクスを記述した確率モデルを観測データから推定し、さまざまな転写制御因子の転写活性を制御する ncRNA を網羅的に同定する。これにより、機能性 ncRNA を効率的に探索するための転写調節ネットワークアトラスを構築することを目的とする(図1)。本研究は、次世代シーケンサーから網羅的に取得された ncRNA を含む大量のトランスクリプトーム情報とスーパーコンピューターを用いた数理モデリングの二つの強力なツールを駆使して、生命科学、情報科学における技術革新と分子生物学を融合する斬新的なアイデアをもとに立案されたものである。

### 3. 研究の方法

本研究では、機能性 ncRNA を効率的に探索するための転写調節ネットワークアトラスを構築することを目的として、以下の研究を行った。

(1). モジュレーター依存的転写制御を記述するための新しい関連の導入

一つまたは複数個のモジュレーター因子の値に応じて変化する転写制御因子とそのターゲット遺伝子群の制御関係を記述するために、エネルギー統計量に基づく局所距離相関(Local Distance Correlation)という新しい相関を定義する。ここで定義された局所距離相関は、一般的に用いられる Pearson や Spearman の相関係数と異なり、1 転写因子と1ターゲット遺伝子の相関のみならず、複数の転写因子群と複数のターゲット遺伝子群との相関を記述することが可能である。また、線形の相関関係以外に、非線形を含む複雑な相関関係を捉えることができる。

(2). モジュレーター因子の網羅的探索のための統計的検定法の開発

モジュレーター因子の網羅的探索のために、(1)で定義した局所距離相関の変化量によって転写因子の転写活性度とモジュレーター因子の関係をモデル化する。次に、転写活性度の変化量を統計量として考え(図2)、その変化量が統計的に有意かどうかをブートストラップ法により検定する統計的手法を開発する。

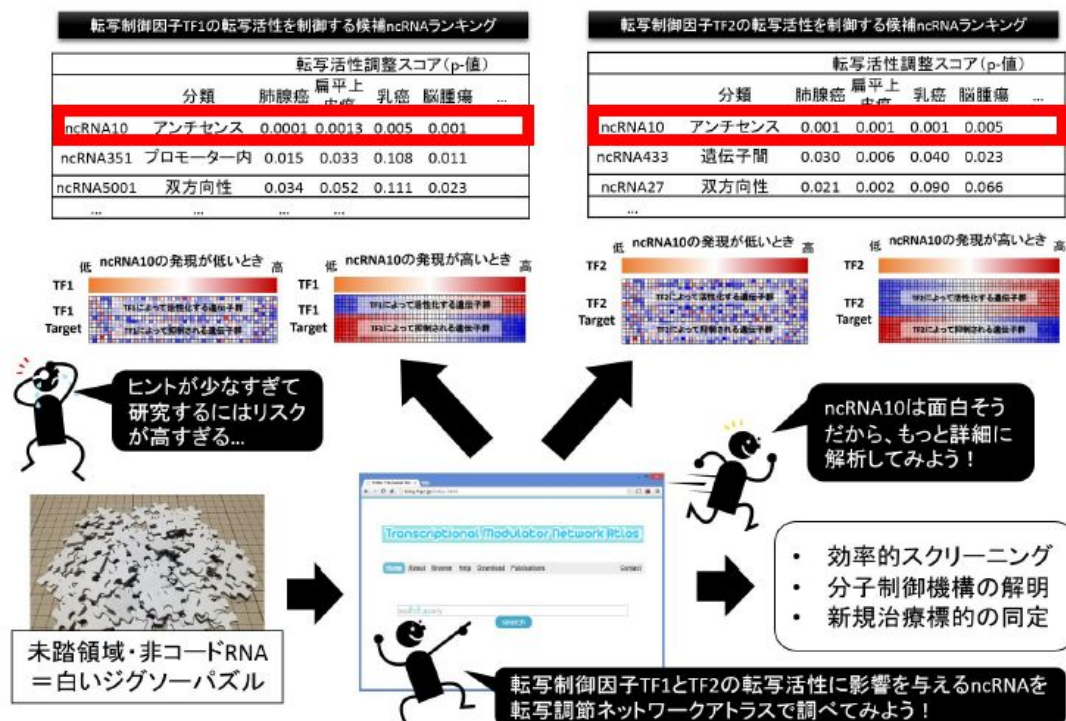


図 1: 転写調節ネットワークアトラスによる機能性 ncRNA の探索

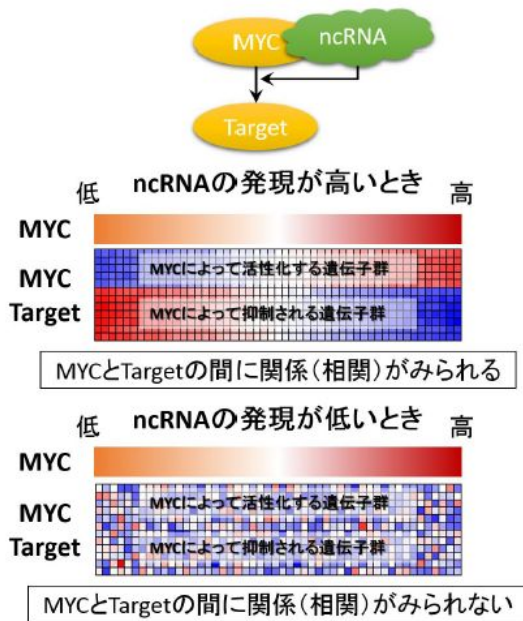


図 2：概念図

(3). ネットワークアトラス構築のための基盤データベースの整備

(1)および(2)で開発する手法を適用する基礎データを整理し、データベース化する。具体的には、大規模ゲノムプロジェクト(ICGC、TCGA、Encode など)および公開データベース(Gene Expression Omnibus、Array Express、Allen Brain Atlas など)に登録されたデータから使用できるデータを精査する。

(4). 機能性 ncRNA の網羅的探索を可能にする大規模パイプラインの構築

(1)および(2)で開発する手法を(3)で得られたデータを基に、全ノンコーディング RNA、全転写因子を対象に、モジュレーター網羅的探索を行うためには、きわめて膨大な計算コストを必要とする。そこで、スーパーコンピュータ上に網羅的解析を行うための大規模パイプラインを構築し、様々な疾患における新規機能性ノンコーディング RNA の同定を可能にする解析システムを構築する。

#### 4. 研究成果

本研究では、モジュレーター依存的転写制御を記述するための局所距離相関、およびモジュレーター因子の網羅的探索のための統計的検定法の有効性の検証のため、シミュレーションデータおよび実データに本開発手法を適用し、以下のような成果を得た。

(1). がん遺伝子 MYC を制御するノンコーディング RNA の発見

1982年に発見された重要ながん遺伝子の1つであるがん遺伝子 MYC の研究は30年以上に研究されてきた。その結果、1万8千報以上の論文が発表されているが、MYC の決定的制御メカニズムは不明のままであった。MYC は数千個のターゲット遺伝子のプロモーターに結合し、転写を制御していることが示唆されていた。MYC と協調して働くモジュレーター因子の多くは未知であり、モジュレーター因子の網羅的探索はがんの分子メカニズムの解明、さらには治療標的となる遺伝子を探る上で極めて重要と考えた。そこで、名大医学系研究科の高橋 隆教授らのグループと連携して、大規模サンプルデータを使って、MYC のモジュレーター因子となりうる長鎖ノンコーディング RNA (lncRNA) を探索するチャレンジを実行した。そのために、エネルギー統計量を用いた局所距離相関に基づく統計的検定によるモジュレーター因子の網羅的探索法(GIMLET)を開発した。GIMLET (Genome-wide Identification of Modulators with Local Energy statistical Test) は、転写因子、ターゲット遺伝子、モジュレーター因子の3項関係を数理モデル化し、観測された遺伝子発現データからモジュレーターの制御調節の変化を推定する。そして、モジュレーターの候補遺伝子の発現が変化したときに、どれくらい転写因子とそのターゲット遺伝子の制御調節を変化させるかを発現データから統計的に検定する。本研究では、東京大学スーパーコンピュータ Shirokane および京コンピュータを用いて、935 サンプルのがん細胞株の遺伝子発現データ、および 596 検体の肺腺癌患者の遺伝子発現データに GIMLET を適用した(図3)。その結果、膨大なノンコーディング RNA の中から、MYC を操っている長鎖ノンコーディング RNA (MYC-modulating lncRNA, MYMLR) を同定した。

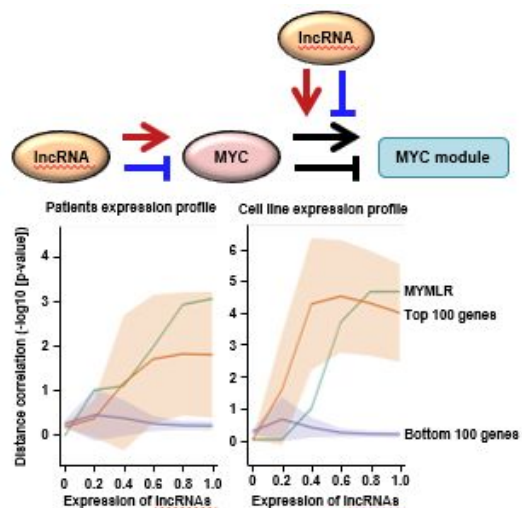


図 3：実データの解析

分子生物学的実験の結果、MYMLR は実際に MYC を制御していることが確かめられ、本解析手法の有効性が確かめられた(現在論

文投稿中)

(2). モジュレーター因子の網羅的探索のための大規模解析パイプラインの開発、および転写調節ネットワークアトラスの構築

遺伝子発現データからのモジュレーターの制御調節の変化量の計算、他のオミクスデータとの統合、モジュレーターの網羅的探索を可能するための並列計算などを行う大規模解析パイプラインを統計言語Rの解析パッケージとして開発した。シミュレーションおよび実データによる検証の結果、本解析パイプラインは従来法よりも大幅に効率的かつ精度よくモジュレーターを探索することができることが確かめられた。具体的には、The Cancer Genome Atlas (TCGA) におけるPan-Kidney コホート研究における 893 症例のマルチオミクスデータに開発した手法を適用し、HIF1A のモジュレーター因子の探索を行った。その結果、これまでに既に報告のある HIF1A モジュレーター因子の同定に成功するとともに、新規のモジュレーター候補を同定できたことから、本解析手法の有効性を検証することができた。また、有効性が検証された手法を TCGA のさまざまながん種のデータに適用し、複数の転写因子に関する機能性ノンコーディング RNA 候補を推定するとともに、その結果をデータベース化した転写調節ネットワークアトラスを構築した。

#### 5. 主な発表論文等

[雑誌論文](計22件)(すべて査読有り)

1. Takeda M, Kanki Y, Masumoto H, Funakoshi S, Hatani T, Fukushima H, Izumi-Taguchi A, Matsui Y, Shimamura T, Yoshida Y, Yamashita JK. Identification of Cardiomyocyte-Fated Progenitors from Human-Induced Pluripotent Stem Cells Marked with CD82. *Cell Rep.* 2018 Jan 9;22(2):546-556. doi: 10.1016/j.celrep.2017.12.057. PubMed PMID: 29320747.
2. Terai H, Kitajima S, Potter DS, Matsui Y, Quiceno LG, Chen T, Kim TJ, Rusan M, Thai TC, Piccioni F, Donovan KA, Kwiatkowski N, Hinohara K, Wei G, Gray NS, Fischer ES, Wong KK, Shimamura T, Letai A, Hammerman PS, Barbie DA. ER Stress Signaling Promotes the Survival of Cancer "Persister Cells" Tolerant to EGFR Tyrosine Kinase Inhibitors. *Cancer Res.* 2018 Feb 15;78(4):1044-1057. doi: 10.1158/0008-5472.CAN-17-1904. Epub 2017 Dec 19. PubMed PMID: 29259014; PubMed Central PMCID: PMC5815936.
3. Park H, Shimamura T, Imoto S, Miyano S. Adaptive NetworkProfiler for Identifying Cancer Characteristic-Specific Gene Regulatory Networks. *J Comput Biol.* 2018 Feb;25(2):130-145. doi: 10.1089/cmb.2017.0120. Epub 2017 Oct 20. PubMed PMID: 29053381.
4. Aoki K, Nakamura H, Suzuki H, Matsuo K, Kataoka K, Shimamura T, Motomura K, Ohka F, Shiina S, Yamamoto T, Nagata Y, Yoshizato T, Mizoguchi M, Abe T, Momii Y, Muragaki Y, Watanabe R, Ito I, Sanada M, Yajima H, Morita N, Takeuchi I, Miyano S, Wakabayashi T, Ogawa S, Natsume A. Prognostic relevance of genetic alterations in diffuse lower-grade gliomas. *Neuro Oncol.* 2018 Jan 10;20(1):66-77. doi: 10.1093/neuonc/nox132. PubMed PMID: 29016839; PubMed Central PMCID: PMC5761527.
5. Taniguchi R, Muramatsu H, Okuno Y, Suzuki K, Obu S, Nakatochi M, Shimamura T, Takahashi Y, Horikoshi Y, Watanabe K, Kojima S. Comprehensive genetic analysis of donor cell derived leukemia with KMT2A rearrangement. *Pediatr Blood Cancer.* 2018 Feb;65(2). doi: 10.1002/pbc.26823. Epub 2017 Sep 17. PubMed PMID: 28921816.
6. Takahashi Y, Sugimachi K, Yamamoto K, Niida A, Shimamura T, Sato T, Watanabe M, Tanaka J, Kudo S, Sugihara K, Hase K, Kusunoki M, Yamada K, Shimada Y, Moriya Y, Suzuki Y, Miyano S, Mori M, Mimori K. Japanese genome-wide association study identifies a significant colorectal cancer susceptibility locus at chromosome 10p14. *Cancer Sci.* 2017 Nov;108(11):2239-2247. doi: 10.1111/cas.13391. Epub 2017 Sep 26. PubMed PMID: 28869801; PubMed Central PMCID: PMC5665761.
7. Tsubota S, Kishida S, Shimamura T, Ohira M, Yamashita S, Cao D, Kiyonari S, Ushijima T, Kadomatsu K. PRC2-Mediated Transcriptomic Alterations at the Embryonic Stage Govern Tumorigenesis and Clinical Outcome in MYCN-Driven Neuroblastoma. *Cancer Res.* 2017 Oct 1;77(19):5259-5271. doi: 10.1158/0008-5472.CAN-16-3144. Epub 2017 Aug 14. PubMed PMID: 28807939.
8. Seki M, Kimura S, Isobe T, Yoshida K, Ueno H, Nakajima-Takagi Y, Wang C, Lin L, Kon A, Suzuki H, Shiozawa Y, Kataoka K, Fujii Y, Shiraishi Y, Chiba K, Tanaka H, Shimamura T, Masuda K,

- Kawamoto H, Ohki K, Kato M, Arakawa Y, Koh K, Hanada R, Moritake H, Akiyama M, Kobayashi R, Deguchi T, Hashii Y, Inamura T, Sato A, Kiyokawa N, Oka A, Hayashi Y, Takagi M, Manabe A, Ohara A, Horibe K, Sanada M, Iwama A, Mano H, Miyano S, Ogawa S, Takita J. Recurrent SPI1 (PU.1) fusions in high-risk pediatric T cell acute lymphoblastic leukemia. *Nat Genet.* 2017 Aug;49(8):1274-1281. doi: 10.1038/ng.3900. Epub 2017 Jul 3. PubMed PMID: 28671687.
9. Shimamura T, Matsui Y, Kajino T, Takahashi T, Miyano S. GIMLET: Identifying biological modulators in context-specific gene regulation using local energy statistics, Proceedings of CIBB2017, 2017 Sep, 1-6.
  10. Matsui Y, Miyano S, Shimamura T, Tumor subclonal progression model for cancer hallmark acquisition, Proceedings of CIBB2017, 2017 Sep, 1-6.
  11. Kondo A, Nonaka A, Shimamura T, Yamamoto S, Yoshida T, Kodama T, Aburatani H, Osawa T. Long Noncoding RNA JHDM1D-AS1 Promotes Tumor Growth by Regulating Angiogenesis in Response to Nutrient Starvation. *Mol Cell Biol.* 2017 Aug 28;37(18). pii: e00125-17. doi: 10.1128/MCB.00125-17. Print 2017 Sep 15. PubMed PMID: 28652266; PubMed Central PMCID: PMC5574049.
  12. Matsui Y, Niida A, Uchi R, Mimori K, Miyano S, Shimamura T. phyC: Clustering cancer evolutionary trees. *PLoS Comput Biol.* 2017 May 1;13(5):e1005509. doi: 10.1371/journal.pcbi.1005509. eCollection 2017 May. PubMed PMID: 28459850; PubMed Central PMCID: PMC5432190.
  13. Kanki Y, Nakaki R, Shimamura T, Matsunaga T, Yamamizu K, Katayama S, Suehiro JI, Osawa T, Aburatani H, Kodama T, Wada Y, Yamashita JK, Minami T. Dynamically and epigenetically coordinated GATA/ETS/SOX transcription factor expression is indispensable for endothelial cell differentiation. *Nucleic Acids Res.* 2017 May 5;45(8):4344-4358. doi: 10.1093/nar/gkx159. PubMed PMID: 28334937; PubMed Central PMCID: PMC5416769.
  14. Kondo A, Yamamoto S, Nakaki R, Shimamura T, Hamakubo T, Sakai J, Kodama T, Yoshida T, Aburatani H, Osawa T. Extracellular Acidic pH Activates the Sterol Regulatory Element-Binding Protein 2 to Promote Tumor Progression. *Cell Rep.* 2017 Feb 28;18(9):2228-2242. doi: 10.1016/j.celrep.2017.02.006. PubMed PMID: 28249167.
  15. Sugimachi K, Matsumura T, Shimamura T, Hirata H, Uchi R, Ueda M, Sakimura S, Iguchi T, Eguchi H, Masuda T, Morita K, Takenaka K, Maehara Y, Mori M, Mimori K. Aberrant Methylation of FOXE1 Contributes to a Poor Prognosis for Patients with Colorectal Cancer. *Ann Surg Oncol.* 2016 Nov;23(12):3948-3955. Epub 2016 Jun 6. PubMed PMID: 27271927.
  16. Matsui Y, Mizuta M, Ito S, Miyano S, Shimamura T. D3M: detection of differential distributions of methylation levels. *Bioinformatics.* 2016 Aug 1;32(15):2248-55. doi: 10.1093/bioinformatics/btw138. Epub 2016 Mar 11. PubMed PMID: 27153581.
  17. Nakaoka HJ, Hara T, Yoshino S, Kanamori A, Matsui Y, Shimamura T, Sato H, Murakami Y, Seiki M, Sakamoto T. NECAB3 Promotes Activation of Hypoxia-inducible factor-1 during Normoxia and Enhances Tumorigenicity of Cancer Cells. *Sci Rep.* 2016 Mar 7;6:22784. doi: 10.1038/srep22784. PubMed PMID: 26948053; PubMed Central PMCID: PMC4780031.
  18. Uchi R, Takahashi Y, Niida A, Shimamura T, Hirata H, Sugimachi K, Sawada G, Iwaya T, Kurashige J, Shinden Y, Iguchi T, Eguchi H, Chiba K, Shiraishi Y, Nagae G, Yoshida K, Nagata Y, Haeno H, Yamamoto H, Ishii H, Doki Y, Iinuma H, Sasaki S, Nagayama S, Yamada K, Yachida S, Kato M, Shibata T, Oki E, Saeki H, Shirabe K, Oda Y, Maehara Y, Komune S, Mori M, Suzuki Y, Yamamoto K, Aburatani H, Ogawa S, Miyano S, Mimori K. Integrated Multiregional Analysis Proposing a New Model of Colorectal Cancer Evolution. *PLoS Genet.* 2016 Feb 18;12(2):e1005778. doi: 10.1371/journal.pgen.1005778. eCollection 2016 Feb. Erratum in: *PLoS Genet.* 2017 May 19;13(5):e1006798. PubMed PMID: 26890883; PubMed Central PMCID: PMC4758664.
  19. Kataoka K, Nagata Y, Kitanaka A, Shiraishi Y, Shimamura T, Yasunaga J, Totoki Y, Chiba K, Sato-Otsubo A, Nagae G, Ishii R, Muto S, Kotani S,

- Watatani Y, Takeda J, Sanada M, Tanaka H, Suzuki H, Sato Y, Shiozawa Y, Yoshizato T, Yoshida K, Makishima H, Iwanaga M, Ma G, Nosaka K, Hishizawa M, Itonaga H, Imaizumi Y, Munakata W, Ogasawara H, Sato T, Sasai K, Muramoto K, Penova M, Kawaguchi T, Nakamura H, Hama N, Shide K, Kubuki Y, Hidaka T, Kameda T, Nakamaki T, Ishiyama K, Miyawaki S, Yoon SS, Tobinai K, Miyazaki Y, Takaori-Kondo A, Matsuda F, Takeuchi K, Nureki O, Aburatani H, Watanabe T, Shibata T, Matsuoka M, Miyano S, Shimoda K, Ogawa S. Integrated molecular analysis of adult T cell leukemia/lymphoma. *Nat Genet.* 2015 Nov;47(11):1304-15. doi: 10.1038/ng.3415. Epub 2015 Oct 5. PubMed PMID: 26437031.
20. Seki M, Nishimura R, Yoshida K, Shimamura T, Shiraishi Y, Sato Y, Kato M, Chiba K, Tanaka H, Hoshino N, Nagae G, Shiozawa Y, Okuno Y, Hosoi H, Tanaka Y, Okita H, Miyachi M, Souzaki R, Taguchi T, Koh K, Hanada R, Kato K, Nomura Y, Akiyama M, Oka A, Igarashi T, Miyano S, Aburatani H, Hayashi Y, Ogawa S, Takita J. Integrated genetic and epigenetic analysis defines novel molecular subgroups in rhabdomyosarcoma. *Nat Commun.* 2015 Jul 3;6:7557. doi: 10.1038/ncomms8557. PubMed PMID: 26138366; PubMed Central PMCID: PMC4506514.
21. Hasegawa T, Mori T, Yamaguchi R, Shimamura T, Miyano S, Imoto S, Akutsu T. Genomic data assimilation using a higher moment filtering technique for restoration of gene regulatory networks. *BMC Syst Biol.* 2015 Mar 13;9:14. doi: 10.1186/s12918-015-0154-2. PubMed PMID: 25890175; PubMed Central PMCID: PMC4371723.
22. Suzuki H, Aoki K, Chiba K, Sato Y, Shiozawa Y, Shiraishi Y, Shimamura T, Niida A, Motomura K, Ohka F, Yamamoto T, Tanahashi K, Ranjit M, Wakabayashi T, Yoshizato T, Kataoka K, Yoshida K, Nagata Y, Sato-Otsubo A, Tanaka H, Sanada M, Kondo Y, Nakamura H, Mizoguchi M, Abe T, Muragaki Y, Watanabe R, Ito I, Miyano S, Natsume A, Ogawa S. Mutational landscape and clonal architecture in grade II and III gliomas. *Nat Genet.* 2015 May;47(5):458-68. doi: 10.1038/ng.3273. Epub 2015 Apr 13. PubMed PMID: 25848751.

〔学会発表〕(計9件)

1. 島村徹平、生命科学におけるビッグデータ駆動型科学の実践と将来、2017年度生命科学系学会合同年次大会(招待講演)、2017/12/6、神戸ポートアイランド、兵庫県。
2. 島村徹平、Analysis of Big Data - Genome Data -, 実世界データ循環学サマースクール 2016(招待講演)、2016/8/4、ハノイ工科大学、ベトナム。
3. 島村徹平、局所距離相関に基づくモジュレーター因子の網羅的探索法、2016年度統計関連学会連合大会(招待講演)、2016/9/6、金沢大学、石川県。
4. 島村徹平、マルチオミックス情報に基づくがんのモジュレーター因子の網羅的探索、第5回生命医薬情報学連合大会(招待講演)、2016/9/30、東京国際交流館ブラザ平成、東京都。
5. 島村徹平、システム生物学的アプローチに基づくオミックス統合解析、第39回分子生物学会年会(招待講演)、2016/11/30、パシフィコ横浜、神奈川県。
6. Shimamura T, Identifying modulators using local energy statistics, *CMStatistics 2016* (ERCIM2016)、2016/12/11、セビリア大学、スペイン。
7. Shimamura T, Discovering modulators of gene regulation by local energy statistics, *IFCS2015*, 2015/7/7、ポロニヤ大学、イタリア。
8. 島村徹平、がん悪性化に寄与する発現制御調節因子のイン・シリコ探索、がんゲノムの情報と数理(招待講演)、2015/9/30、東京大学医科学研究所、東京都。
9. 島村徹平、がんの俯瞰的理解のための統合的統計モデリング、第38回日本分子生物学会年会(招待講演)、2015/12/2、神戸ポートアイランド、兵庫県。

〔図書〕(計1件)

1. 松井佑介、島村徹平(2018) *phyC - がん進化を推定・分類するためのデータ駆動型数理アプローチ: 遺伝子医学 MOOK 33号「遺伝統計学と疾患ゲノムデータ解析 - 病態解明から個別化医療, ゲノム創薬まで -」*

〔その他〕

ホームページ等

<http://www.nagoya-sysbiol.info/>

6. 研究組織

(1)研究代表者

島村 徹平 (Shimamura, Teppei)

名古屋大学・大学院医学系研究科・特任准教授

研究者番号: 00623943