

システム癌新次元Newsletter No. 1

2016年3月

がんシステムの新次元俯瞰と攻略

文部科学省科学研究費補助金 新学術領域研究 (研究領域提案型)(複合領域:4701)

システム癌新次元

ー がん研究は人智を超えた領域に突入

領域代表者 宮野 悟

東京大学医科学研究所ヒトゲノム解析センター



新学術領域「システム癌新次元」（複合領域：4701）は、2010年から5年間展開してきました新学術領域「システムがん」（複合領域：4201；2010年度～2014年度）を深化させることを目的に提案いたしました。2009年当時、私はがんに限らず、生命をシステムとして理解するため、スパコンを使った計算システム生物学の研究に没頭していました。そのときがん研究のトップランナーの方々に「これまで我々はがん関連遺伝子のジャングルを這い回って研究していた。これを続けても限界だ。」といわれ、数学とスパコンを駆使した大規模データ解析と数理モデリングをがん研究に融合する「システムがん」の着想を得ました。その結果、がんのシステム異常の俯瞰的理解が飛躍的進み、がん研究の歴史に刻まれる発見もありました。また、同種のがんだけでなく、一患者さんにおいても想像を超えた生体空間時間的多様性が明らかになるにつれ、美しいがん生物学の終焉が見えてきました。がんの理解は、もはや生物学・医学の領域から昇華し、新しい次元へのオデッセイが始まったと考えています。

私達が経験したことのないデータ量の世界に突入し、ゲノムデータだけでも既に数百ペタバイトになっています。がんの全ゲノム解析をすると数千～数百万の変異が見つかってきますが、その生物学的・臨床的解釈は難しく、途方に暮れてしまう毎日でした。これまでは、エクソンという全

ゲノムの1.5%内外しか調べていなかったことがほとんどですが、全ゲノムシーケンスによって、エクソン以外の複雑な異常が、遺伝子の機能に決定的な変更を加えてしまうこともわかってきました。機能を持ったノンコーディングRNAの未踏の地の出現は、エクソンだけをみてきた者にとって「我々は井の中の蛙だった」ことを知らしめさせられました。さらに、がんに関する論文は2014年だけで20万報を超え、急増しています。文献検索をして論文を読む、データベース検索をするといったこれまでの方法が通用しなくなっています。まさに人智を超えた領域に踏み込んでいます。

また、Germline、すなわち胚細胞に刻まれた遺伝情報、言いかえますと、皆さんが、親子で継承するゲノムDNAには、バリエーション（変異）の存在が良く知られています。そのバリエーションは、一人一人の個性を生み出している源であることもご存知の通りです。しかし今、何万人もの全ゲノムシーケンスデータが蓄えられその情報解析が進んできました。その結果、この「胚性バリエーションの効果」は予想より遙かに大きいという、「不都合な事実」に直面することになりました。その象徴的な出来事は、女優で映画監督のアンジェリーナ・ジョリーさんの予防的乳房摘出手術、さらに最近では卵巣・卵管摘出手術の報道でした。わずか125年前の明治時代、平均寿命は50才（M25: 男43、女44）以下でした。

しかし、男女とも80才近くになった今、この「胚性バリエーション」は人の生涯において「個性」という言葉では見過ごせない極めて大きな影響を与えることになりました。また、0才～80才までの50,000人以上を対象にしたゲノムバリエーション解析から、実は、私たちの身体の組織や臓器を構成している正常な細胞で、いつの間にか「新たにクローナルな変異が生じている」という事実が明らかになりました。50歳をすぎたころからがんはあなたに巣くいはじめるといってもよいでしょう。ゲノムは一生変わらないというのは正しくなくなったわけです。私たちは年齢を重ねていくことから逃れることはできません。論文誌Cellに2013に発表された“The Hallmarks of Aging”という総説には、9個の「老化の特徴」が示されており、この「衰え」の細胞特性は、がんの病態の特性そのものでした。さらに、これら「9つの特徴」はがん以外の生活習慣病の細胞機能の変化そのものでもあり、細胞文脈(cellular context)という変調を生み出しています。細胞文脈は、ゲノムの情報だけでこれを理解できるものではありません。細胞文脈の理解には、遺伝子と蛋白、そしてその機能ネットワークという3次元でとらえられるものでなく、そこには時間軸、即ち加齢という尺度を加えた新次元のアプローチが必要となっています。

私たちはこのような「不都合な事実」に立ち向かわねばならない状況になったわけです。先

の5年間の「システムがん」の取り組みの中で、私たちは、この「不都合な事実」を目の当たりにし、これに対応すべく、倫理、社会的な取り組み即ち「倫理的法的社会的課題 (ELSI: Ethical, Legal and Social Issues)」の重要性を強く受け止めました。しかも、この「ELSI」の取り組みは、基礎研究とリアルタイムで、両輪で推進されなければ意味がないと認識するに至りました。がんの新発見に直面する未遭遇の高齢化社会が登場し、社会学が深く関わる時代が始まりました。

2014年度に終了した「システムがん」では、数理的方法論とスパコンで大規模データを詳細に俯瞰することはできるようになりました。しかし、数千～数万の多数の試料検体データ、ノンコーディングRNAを含む数万遺伝子からなるネットワークは、ただただ複雑です。

全ゲノムシーケンスの結果、機能が不明である膨大な数の変異が現れてきました。しかし、現在、文献を調べようにも、Medlineには2000万報以上。報告には間違いもあります。COSMIC (Catalogue of Somatic Mutations in Cancer) に報告されている遺伝子変異だけでも「目」と「手」で調べることは無理です。システムがんの成功は、私の目からみますと、班員研究者の超人的能力と努力により、複雑・大量の俯瞰情報から「多くの知見と深い洞察力」により初めて実現されたものと思います。

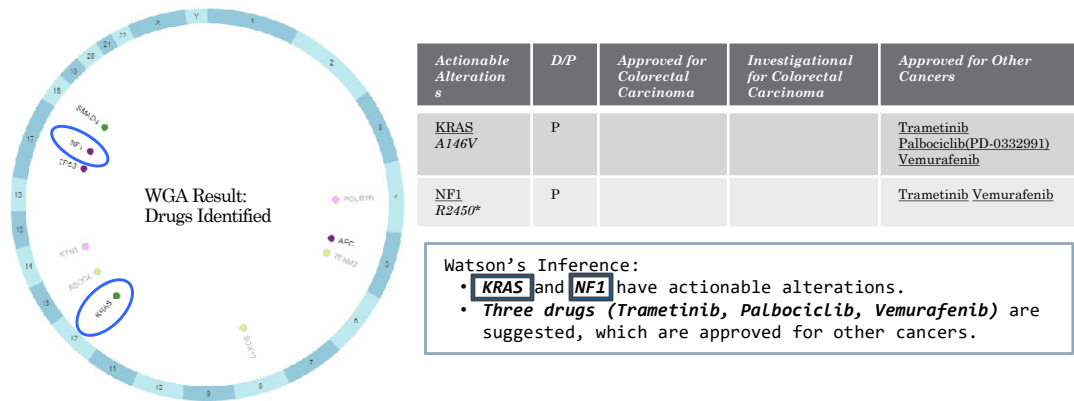


図 1

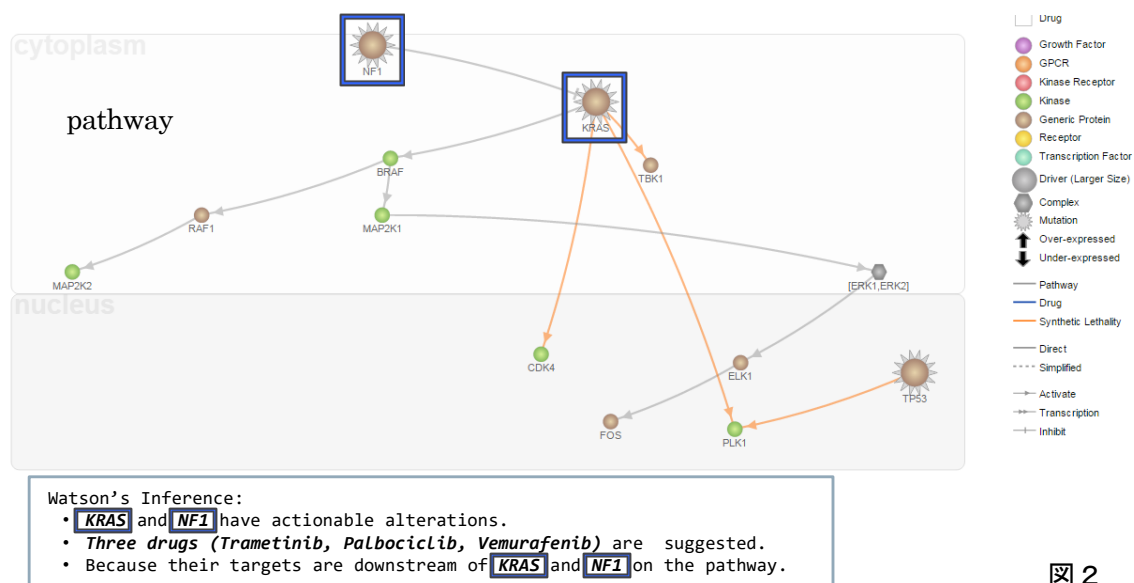


図 2

「システム癌新次元」の第一の戦略は、「システムがん」を推進させた基盤ソフトウェアを新たなスーパーコンピュータで発展させること、第二の戦略は人智を超える方法論を導入することです。東京大学医科学研究所ヒトゲノム解析センターには2015年4月からShirokane3というスーパーコンピュータが稼働をはじめました。計算能力は414TFLOPS、Lustreとよばれる高速ディスクアレイは12PBあり、100PBまで拡張可能なストレージが用意されています。さらに、理化学研究所計算科学研究機構の「京」コンピュータも利用でき、さらに現在開発が進んでいるポスト「京」を利用可能だと考えています。第二の戦略の一つとして、東京大学医科学研究所に2015年7月から導入されたIBM Watson Genomic Analytics (WGA)があります。Watsonは自然言語（英語）を理解し、学習し、推論し仮説を重みつけて提示する、学習するデータベースのようなものでCognitive Computingとよばれる技術が使われています。導入時の学習内容（New York Genome Centerで訓練）は、2000万件超のMedlineデータ（文献アブストラクト）、1500万件超の薬の特許データ、COSMIC (Catalogue of Somatic Mutations in Cancer,

UK)というがんにおける体細胞変異データベース、ClinVar (Genomic Variationとhealthに関する情報, NIH USA)、National Cancer Institute Pathways (NIH USA)などのデータを学習しています。WGAの雰囲気だけですがお伝えします。試料は大腸がんの細胞株を用いています。エクソーム解析で体細胞変異が230出てきました。この変異リストをWGAにアップロードすると10分程度で図1のようなドライバー遺伝子の候補と分子標的薬の候補（選ばれてきた理由も含め）が得られます。また図2のようにパスウェイによる説明をでできます。全ゲノムシーケンスをして100万ほどの変異候補をいれても30分程度でドライバー遺伝子と薬が、文献やデータベース情報による理由とともに提示されます。人智を超えられるかもしれないという期待感が現在高まっています。

このような目標を掲げ、6人の研究者が「システム癌新次元」を構想しました（図3）。それぞれの学問的バックグラウンドは、稲澤譲治（血液内科）、高橋 隆（外科）、小川誠司（血液内科）、岡田随象（ゆきのりと読みます。リウマチ）、武藤香織（社会学）、宮野 悟（数学）と多様です。

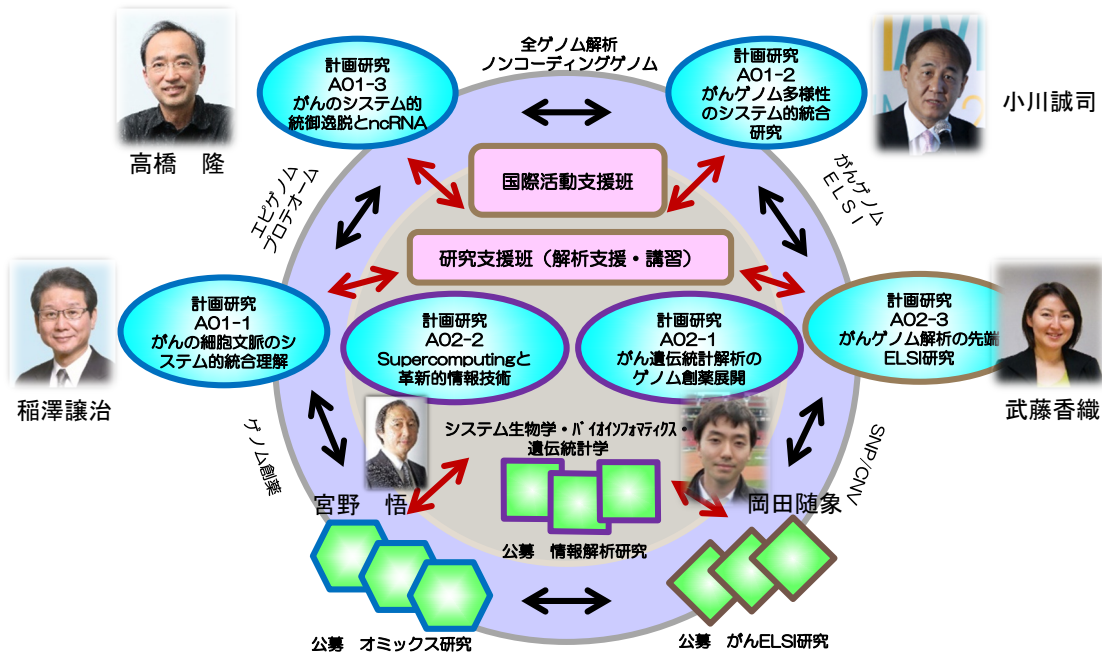
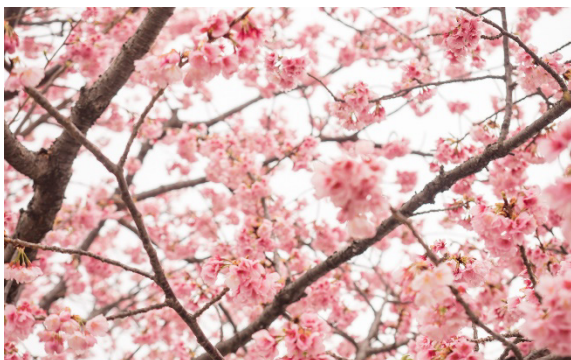


図3 システム癌新次元連携体制

この新学術領域は「システム癌新次元」と称しておりますが、私達が創ろうとしているものは、「がん」という比較的私達班員がよくわかっているつもりになっているテーマをとりあげ、スーパーコンピュータ「京」やポスト「京」を駆使し、大きな可能性を秘めたビッグデータ、さらには高度に発達した人工知能システムが引き起こす、私達が遭遇したことのない

新次元研究の端緒となることを目指したものです。私どもは、がんにかかわらず、生命科学、さらには、他の分野における、深い基礎を作ることができる学問パラダイムの構築を第一の目的にしています。

「システム癌新次元」の領域を代表して
東京大学医科学研究所
宮野 悟



計画研究班の構成

A01 がんのシステムの統合理解の新展開

A01-1 がん細胞文脈のシステムの統合理解による新たながん診断・治療概念の確立

研究代表者 稲澤 譲治 東京医科歯科大学難治疾患研究所 教授
研究分担者 井上 純 東京医科歯科大学難治疾患研究所 講師
研究分担者 村松 智輝 東京医科歯科大学難治疾患研究所 助教
研究分担者 谷本 幸介 東京医科歯科大学難治疾患研究所 助教
連携研究者 石川 俊平 東京医科歯科大学難治疾患研究所 教授

A01-2 大規模シーケンスによるがんの遺伝学的・生物学的多様性のシステムの統合研究

研究代表者 小川 誠司 京都大学大学院医学研究科 教授
連携研究者 吉田 健一 京都大学大学院医学研究科 助教
連携研究者 片岡 圭輔 京都大学大学院医学研究科特定助教

A01-3 肺がんの分子病態をノンコーディングRNAから俯瞰するシステムの統合研究

研究代表者 高橋 隆 名古屋大学医学研究科 教授
研究分担者 梶野 泰祐 名古屋大学医学研究科 特任助教
研究分担者 柳澤 聖 名古屋大学医学研究科 講師
研究分担者 中枳 昌弘 名大病院 病院助教
研究分担者 長田 啓隆 愛知県がんセンター研究所 室長

A02 がんビッグデータ — 情報解析の革新とELSI

A02-1 遺伝統計学によるがんシステムのモデル化とゲノム創薬手法の開発

研究代表者 岡田 随象
東京医科歯科大学大学院医歯学総合研究科 テニユアトラック講師
研究分担者 藤本 明洋
理研統合生命科学研究センター チームリーダー
研究分担者 鎌谷 洋一郎
理研統合生命科学研究センター 上級研究員
研究分担者 浦山 ケビン 東京医歯大院 テニユアトラック講師

A02-2 スーパーコンピューティングと人工知能応用技術によるがんシステムの新たな次元探索

研究代表者 宮野 悟 東京大学医科学研究所 教授
研究分担者 井元 清哉 東京大学医科学研究所 教授
研究分担者 山口 類 東京大学医科学研究所 准教授
研究分担者 白石 友一 東京大学医科学研究所 助教
研究分担者 島村 徹平 名大院医学系 特任准教授
研究分担者 玉田 嘉紀 東大院情報理工 助教
連携研究者 樋口 知之 統計数理研究所 所長

A02-3 ゲノム解析の革新に対応した患者中心主義ELSI（倫理・法・社会的課題）研究の構築

研究代表者 武藤 香織 東京大学医科学研究所 教授
研究分担者 丸 祐一 鳥取大地域学 准教授
研究分担者 東島 仁 信州大医学系 特定雇用助教
研究分担者 井上 悠輔 東京大学医科学研究所 助教

外部諮問委員

中村 祐輔 シカゴ大学教授
北川 源四郎 大学共同利用機関法人情報・システム研究機構長
鎌谷 直之 株式会社スタージェン会長

Information

オフィシャルWebサイト

<http://neosystemscancer.hgc.jp/>

システムがんの研究内容、構成研究者情報、発表論文、研究集会などの詳しい情報はこのWebサイトをご覧ください。

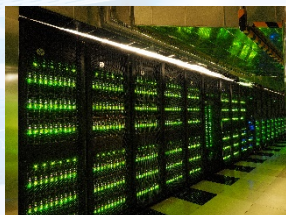


Human Genome Center

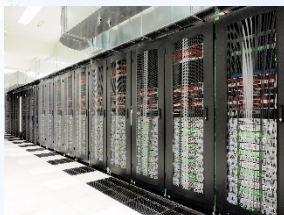
東京大学医科学研究所
ヒトゲノム解析センター
スーパーコンピュータ

<http://supcom.hgc.jp>

システム癌新次元で用いられるデータ解析ソフトウェアは、ヒトゲノム解析センターのスーパーコンピュータで動いています。スーパーコンピュータはどなたでもご利用になれます（有償）。



大容量のディスク装置



高密度の計算機

新学術領域研究「システム癌新次元」ニュースレター No. 1

発行日★平成28年3月17日（初版）

発行★がんシステムの新たな俯瞰と攻略

領域代表者・宮野 悟

東京大学医科学研究所 ヒトゲノム解析センター DNA情報解析分野

〒108-8639 東京都港区白金台4-6-1

TEL: 03-5449-5615 FAX: 03-5449-5442

E-mail: miyanolab-jimu@edelweiss.hgc.jp