

システムがんNewsletter No.4

2012年3月

システムの統合理解に基づくがんの先端的診断、治療、予防法の開発

文部科学省科学研究費補助金 新学術領域研究（複合領域：4201）

研究の紹介

世界で進む網羅的がん細胞ゲノム配列の解読と解析、そして個別化医療へ

角田 達彦（研究代表者）

理化学研究所ゲノム医科学研究センター統計解析・技術開発グループ グループディレクター



正常細胞ががん細胞に変わる一番根源的なものはゲノムDNA配列の変化です。一人の人が一細胞ごとに持つゲノムDNA配列は62億塩基対という膨大なデータからなり、ヒトゲノム計画では10年越しで解読しました。しかし、今や解読する機械であるシーケンサーの発達が目覚ましく、いわゆる次世代シーケンサーを使いますと、平均して半日もあれば、一人の全ゲノムを十分高精度に読み出すことができるようになりました。とはいえそのような機械から出てくるデータ自体は、まるでシュレッダーにかけた紙のようなもので、だいたい50~100文字ぐらいのものばかりで（一つ一つをリードといいます）、それらをパズルのようにつなぎ合わせる必要があります（次ページの図をご参照ください）。また、リードにはちらちら読み間違いが入っているために、同じ場所でも何十回も読んで多数決をとることに似た方法をとって、誤りを無くさないといけません。そして、がん細胞のゲノムDNA配列は、同じ人の正常細胞のゲノムDNA配列と異なる箇所がたくさん見られ、それらの中に、がんを引き起こす犯人がいると考えられます。そこで、がん細胞と正常細胞との間のゲノムDNA配列の違い、すなわち2つのゲノムDNA配列の違いを正確に見いだす必要があります。そこで私たちは、次世代シーケンサーを使って一人の日本人の全ゲノムを読み取り、例えばベジアンと呼ばれる数学的手法などを駆使し、また検証実験を重

ねることによって、高精度、かつスーパーコンピュータを使って高速に解析する方法を確立していきました。その解析では、一文字の違い、配列の欠失や挿入やその数の違い、逆転する部分、位置が飛んでいる部分や、これまで見つかってこなかった配列などを全てあらいだすことができます。その成果を2010年に Nature Genetics 誌より報告しました。ちなみに日本人の全ゲノムを解読し包括的に解析したのはこれが世界で初めてということで、大変話題になりました。

さて、そのようにして次世代シーケンサーの技術を用いて高速・高精度にがんのゲノム上の変化をリストアップすることができるようになってきたわけですが、2008年4月、50種類ほどのがんのゲノム上の変化を網羅的にカタログ化することを目的とする国際がんゲノムコンソーシアム (International Cancer Genome Consortium; ICGC) が発足しました。一つのがん種ごとに500人の患者からがん細胞と正常細胞の両方のDNAをインフォームドコンセントを得た上で頂き、正常とがんを合わせて50,000ゲノムをシーケンスし、25,000がんゲノム上の変化のリストアップします。それらを例えばがん種ごとにまとめ、そのがんの原因となっている遺伝子、原因となる遺伝子群の経路（パスウェイといいます）や、変化のパタンの特徴などをつきとめていきます。

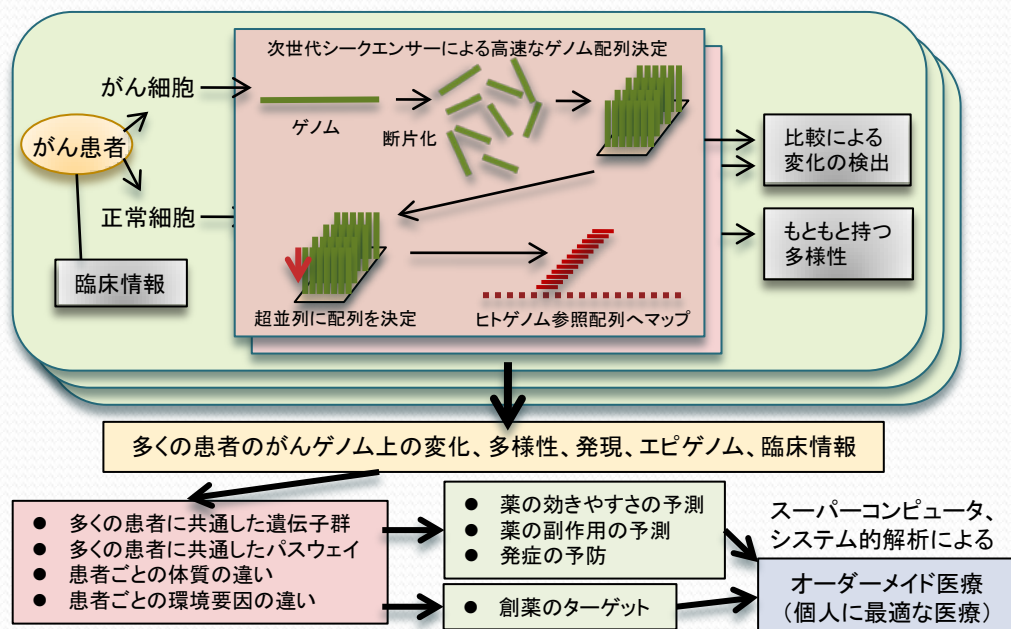
（次ページへ続く）

そしてそれらの遺伝子やメカニズムがわかるにつれ、どこを治療のターゲットとすべきかなどがわかってきます。それをもとに、新たな薬の開発を行ったり、既存の治療薬の中から個々人のがんゲノムの変化のパターンに応じて最もふさわしいものを選んだりするといったオーダーメイド医療が可能になります。このICGCに日本から、私達すなわち理化学研究所ゲノム医科学研究センターと国立がん研究センターが参加し、主にウイルス性の肝がんの解析を行っています。前述の高精度な解析方法の確立とともに、次世代シーケンサーの技術のさらなる進展やコストダウンにより、多症例のシーケンスが進み、私達日本のグループをはじめ、英米中などからも多くの成果を出し始めました。がん細胞のゲノム上での変化のカatalog作りのみならずがんを引き起こすパスウェイなどが次々と上がってきています。私達日本のグループも、肝がん27症例の詳細な解析結果を論文化しています。この肝がん全ゲノムの多症例の解析および前述の日本人ゲノムの解析は、東京大学医科学研究所ヒトゲノム解析センターのスーパーコンピュータを使い、宮野悟先生とその研究室の方々のご協力のもと行うことができました。

つい先日、2012年3月20から22日まで、フランスのカヌで、このICGCの第6回目のワークショップが開

催されました(今回は京都で私達が主催しました)。このワークショップでのトピックは、コンソーシアムの進捗状況の報告、目立った進展のあったグループによる発表(日本も私達のグループより中川英刀が肝がんゲノムの解析結果について発表)、実験による検証方法についての議論、新しいシーケンサーやエピゲノム・プロテオミクス技術の動向の報告、各国のポスター発表、倫理面やデータアクセスの検討、バイオインフォマティクス、臨床への橋渡しなどでした。最初の進捗報告では、現在16カ国が参加し、47プロジェクトに達し、既に3,561のがんのゲノム上の変化がCatalogされたと報告がありました(正常細胞のゲノムも合わせると6,000以上)。また膨大なデータ解析が必要であり、コンピュータ資源の共有が課題になってきていることや、検証実験による高精度な解析の保持をすること、そしてアメリカ The Cancer Genome Atlas (TCGA) というもう一つの大きなプロジェクトなどとの統合解析について議論が行われました。実験による検証に関しては、正解となるデータセットを構築する話に多くの時間がさかれました。少し詳しく述べますと、次世代シーケンサーがときどき起こす読み間違いを防ぐには、同じ場所を多数回読んで信頼性を増す方法をとることを前述しましたが、実際には、例えば40回読んで数回

ゲノムDNAを中心とした大規模オミックスデータの情報解析と遺伝統計学的解析によって可能になる、安定で確実な予測に基づくがん研究と個人の医療診断



ほど異なる文字になったとき、それが本物で(もともと父方と母方からのゲノムの両方を同時に読んでるので、異なる2文字が重なることがある)偶然それらが少なく読まれたのか、単なる読み間違いが重なったのか、なかなか判断が難しいときがあります。そのようなあやしいものは全て間違いとして捨てると保守的すぎ、本物を見落とす可能性もありますのでさじ加減になるわけです。その線引きを、より確実な(だが時間がかかる)実験で検証し、正解セットをできるだけたくさん作って決めることが必要になります。そのための正解セットの構築をICGCとして行う計画が進められています。また新しい技術の動向に関しては、がんゲノムDNAは最も大事な出発点ですが、それから転写される遺伝子の発現プロファイル(トランスクリプトーム)や、がんゲノムDNA上のエピゲノムといわれる修飾情報など、同じ次世代シーケンサーを工夫して使うことによって得られるデータも多く、それらをいかにして統合して解析していくかが、最も重要な課題となりつつあり、この「システムがん」の考え方に大変沿っています。バイオインフォマティクスのセッションでは、それらのいろいろな階層をヒトゲノム上の一部の領域で重ねて統合解析したENCODEプロジェクトの全容紹介や、さまざまながんでの遺伝子内に限る変異をカタログ化しデータベースしつつあるCOSMICからの解析例、そしてがんの真の原因をいかにして見つけ出すかという方法論が論じられました。そして臨床への橋渡しのセッションでは、これらのデータを統合して、例えば乳がんなどを細かく分類し直し、個人ごとの治療法の選択を行うというデザインとデータが示されました。またがん細胞と一口に言っても、がんの進行に応じてゲノム配列が変化していきます。とくにがん化の初期の段階と、進行の進んだ段階とで、原因となっている変化を見極めることによって、より詳細なメカニズムの理解と、治療法の選択につながります。また、がんの中心となっている細胞と、周囲の細胞とのゲノム配列の違いを見極めることも同様に大事で、これらへのアプローチがいくつか報告されました。

以上、がんゲノム上の変化のカタログを構築する国際がんゲノムコンソーシアムの動向と、次世代シーケンサーによるがんゲノムの解読と解析を中心として話をしてみました。抗がん剤の効果はがん細胞の増殖をおさえ消失させることですので、がん細胞の性質をよく見極め、適切な抗がん剤を用いる必要があります。それに対し抗がん剤の副作用は、正常細胞側の負担になって現れるものです。そこで、その人のもともと持っている体質、ひいては生まれついて持っているゲノムDNA配列などを調べ、副作用の少ない薬を選び必要があります。またがんへのなりやすさなりにくさにも個人差があり、それがどの遺伝子やメカニズムによるものなのかを調べておくことも大変重要です。これらに対して私達のグループは、多くのがん患者さんの血液(正常細胞)のDNAと一般の方々のDNAとを見比べ、がんに関連する遺伝子を網羅的に探索するゲノムワイド関連解析という解析も行っており、肝がんウイルスを持っている人の肝臓のがん化、前立腺がんへのなりやすさ、乳がんのタモキシフェン治療の効果、大腸がんへのなりやすさをはじめ、多数のがんに関わる遺伝子の発見を行ってきました。その解析方法は主に一文字単位の個人差を対象にしますが、欠失や挿入などのより大きな単位・複雑な単位も解析できるような手法の開発も行ってきました。また複数の遺伝子が相互作用を起こし合っただん発症のリスクを上げる場合にあり、それを発見する高速な手法を開発し、スーパーコンピュータ上で実行可能にしました。そしてそれらに加え食事や生活習慣などの環境要因なども関わってきますので、臨床応用のためには、がん全体と、個人全体も同時に考え、予防や治療法の選択が可能となる予測を行う方法が必要となります。それらはネットワーク構造になっていますので、膨大なデータの取得や適切なモデル化と検証を行うために、体系的なアプローチを用いたスーパーコンピューティング手法の開発、そしてそのための人材育成が急務となっております。

乳がんの分子サブタイプ分類と個別化抗がん剤治療の開発

三木 義男¹（研究代表者）
松浦 正明²

1 東京医科歯科大学難治疾患研究所分子遺伝分野
2 （公財）がん研究会がん研究所がんゲノム研究部



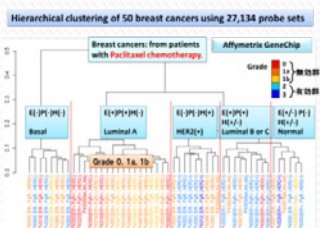
抗がん剤治療では治療効果や副作用にしばしば個人差がみられます。このようながん治療の個人差を規定する因子として、(1) がん組織そのものの個性（がんの個性）、(2) がん患者自身の遺伝要因（患者の体質）が挙げられます。近年、ゲノム情報や遺伝子発現情報を網羅的に解析する新しいオミックス技術が開発され、これらの技術を駆使することによって多数の症例を効率的に解析し、それぞれの臨床病理学的・薬理学的特性と対比することにより「がんの個性」や「患者の体質」を規定する遺伝情報を系統的にスクリーニングすることが可能になりました。そ

こで、私たちの研究グループでは、乳がん患者さんを対象に、抗がん剤投与前のがん組織における網羅的遺伝子発現情報を応用し、乳がんの抗がん剤感受性の予測システムを検討してきました。しかし、網羅的解析は複数の予測モデルを生みだし、加えて乳がんの高度な多様性のためその中から高精度のシステムを同定することは容易ではありません。事実、私たちが作製したパクリタキセルという抗がん剤の感受性予測診断システムも、検証のための臨床試験では奏効率の改善は見られず、乳がんの高度な多様性に対する対応の再考を迫られていました。

先進的遺伝子発現情報に基づく乳癌の多様性解析とその分子病態の理解および個別化治療への展開

乳がんの高度な多様性

1. 乳がん発生母地（腺管の種類）
 - luminal epithelium
 - basal epithelium
2. ホルモン環境
3. 遺伝的多様性
4. 分子サブタイプ



この多様性の理解が最適な診断・治療法の開発には必須！

階層的情報である「網羅的 mRNA情報」、「miRNA情報」、「ゲノム一次構造情報」の統合的利用システムを開発し、多様性を理解！



高度な情報科学的解析技術を駆使

- ① 乳がんサブタイプの簡便な分類法を構築
- ② 分子サブタイプの生物学的機能の特徴を発現情報、miRNA・ゲノム構造情報から描き出し、病態の違いを細胞機能システムとして理解

- ① 分子サブタイプ毎に抗がん剤感受性予測システムを整備→個別化抗がん剤治療の実現
- ② 乳がんの根幹に流れる分子病態を導き出す可能性

Sorlieらは、遺伝子発現情報に基づいたクラスター解析という方法で、乳がんを5~6種のサブタイプ (intrinsic subtype) に分類し、臨床的転帰にも特徴を持つ分類であることを示しました(Sorlie et al. PNAS, 2001)。現在、乳がん治療を選択するうえで、このサブタイプ情報を考慮することが非常に重要とされています。そこで私たちは、パクリタキセルによる抗がん剤治療を受けた乳がん患者さんを、遺伝子発現情報を基に5種のサブタイプに分類し、パクリタキセル治療効果と組み合わせたと、luminal A というサブタイプに属する症例はすべてパクリタキセルが無効でした。この結果より、乳がんは多様であるため、サブタイプ毎に有効例の頻度が異なり、サブタイプ毎に感受性予測システムを作製する必要があると考えました。そこで、私たちは、乳がんの網羅的遺伝子発現情報に、新たに microRNA (miRNA) 情報、ゲノム構造情報を加え、高度な情報科学的解析技術を駆使して、(1) クラスタリング法に代わるintrinsic subtype の簡便な分類法を考案し、分子サブタイプ毎に抗がん剤感受性予測システムを整備することによって個別化抗がん剤治療の実現を目指しています。また、(2) 乳がんサブタイプの生物学的機能の特徴を遺伝子発現情報、miRNA・ゲノム構造情報から描き出し、病態の違いを細胞機能システムとして理解しようとしています。その結果、乳がんの根幹に流れる分子病態を導き出せる可能性があります。



リン酸化プロテオミクスを用いた がん細胞特異的シグナル伝達機構の解明

杉山 直幸¹（研究代表者）

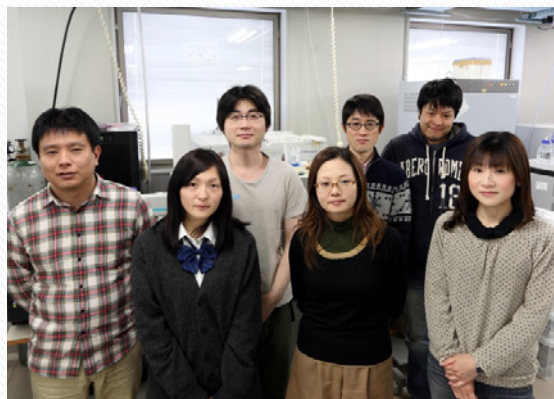
石濱 泰²

¹ 慶應義塾大学先端生命科学研究所

² 京都大学大学院薬学研究所

生体内では、遺伝子やタンパク質といった様々な分子が、生命活動を行うためにそれぞれの役割を果たしています。細胞内の分子が、状況の変化に対応して正しく機能するためには、その情報を個々の分子へと伝達する必要があります、これを細胞内シグナル伝達と呼びます。がん細胞ではシグナル伝達が正常に行われない為に、細胞分裂の亢進やアポトーシス（細胞死）の抑制が異常に引き起こされています。

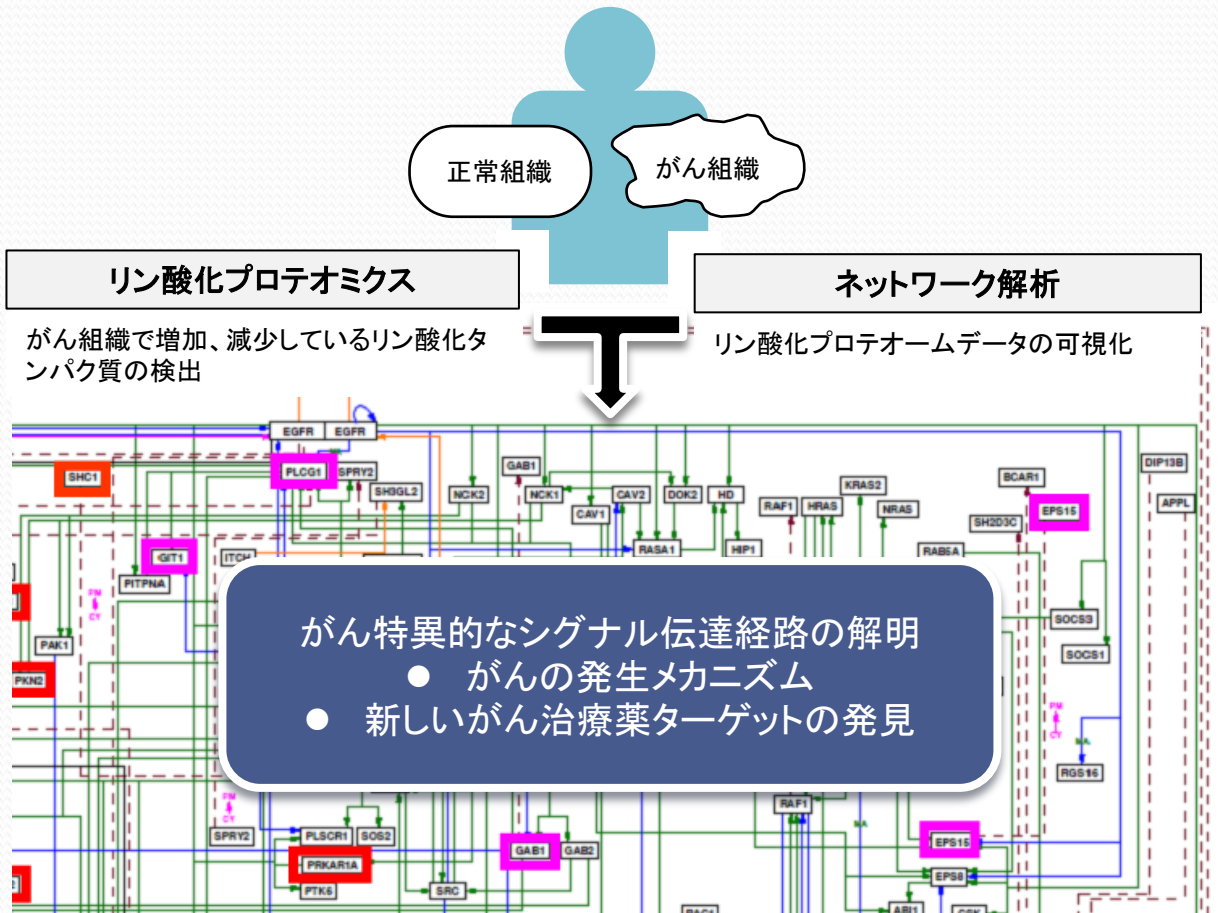
タンパク質のリン酸化修飾はシグナル伝達において重要な因子の一つです。タンパク質は、プロテインキナーゼという酵素によってリン酸化されることにより、機能や構造が変化します。その結果、今度は別のタンパク質がリン酸化されるといったリレーが繰り返されることで、シグナルを他の分子へと伝達していきます。がん細胞では、チロシンキナーゼ型受容



左端が研究代表者（杉山）

体の過剰発現や調節機構の損失などにより、正常細胞と比較して細胞分裂に関わる様々なシグナル伝達経路が活発化しており、がん治療薬の開発において、プロテインキナーゼはその標的分子として、近年非常に注目されています。一方で、シグナル伝達の経路は非常に複雑なネットワークを成しており（図を参照）、がんにおいても関与する分子や伝達経路の全貌は未だに明らかになっていません。





私達の研究グループでは、タンパク質のリン酸化を大規模に解析する”リン酸化プロテオーム解析”を用いて、がん組織に特異的、または過剰に亢進しているシグナル伝達経路や関与するプロテインキナーゼを明らかにすることを目的とした研究を行っています。キーテクノロジーとして、私達はこれまでにリン酸化ペプチド濃縮法 (Hydroxy Acid Modified Metal Oxide Chromatography: HAMMOC法) の開発と、その手法を用いたリン酸化プロテオーム解析技術の確立を行いました。培養細胞を試料としたリン酸化プロテオーム解析結果からリン酸化情報の蓄積を行い、ヒトタンパク質において現在90,000箇所以上のリン酸化部位を同定しています。現在は、がん組織と正常

組織それぞれから抽出した試料についてリン酸化プロテオーム解析で比較することで、がん組織特異的にリン酸化が亢進しているタンパク質の解析を進めています。前述のようにシグナル伝達経路は複雑かつ不明な経路も多いことから、データの可視化法とキナーゼ基質の予測についても評価を進めています。

将来的には、ゲノム、トランスクリプトーム、メタボロミクスと合わせたマルチオミクス解析により、詳細ながんの理解が可能になるものと考えます。がんの発生メカニズム解明や、診断や予防、治療薬開発において新しい創薬ターゲットの発見につながることを期待して、以上の研究を進めております。



文部科学省新学術領域研究

システムがん

システムの統合理解に基づくがん
の先端診断、治療、予防法の開発

オフィシャルWebサイト

<http://systemscancer.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。



<http://twitter.com/SystemsCancer>

システムがんの研究成果、世界のがん研究、研究拠点、学会情報など最先端な話題をつぶやいています。

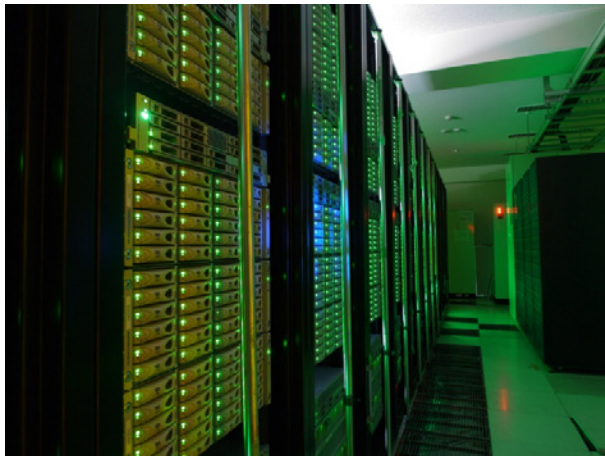


Human Genome Center

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/>

システムがん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できます(有償)。



大容量のディスク装置



高密度の計算機

編集後記

遅れていた桜の開花に合わせてように発行することができましたニュースレターの第四号です。「システムがん」はいよいよ後半の期間に入ります。班員の皆様に対しましては、より多くのご面倒をいただくことになってしまいますが、変わらずのご協力をいただけますようよろしくお願いいたします。

新学術領域研究「システムがん」ニュースレター No.4

発行日★平成24年3月31日(初版)、平成24年4月2日(第2版)

発行★システムの統合理解に基づくがんの先端診断、治療、予防法の開発
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

編集★サトウアユ