



記念講演(1)

ヒトゲノム計画の進展と21世紀の生命科学

榊 佳之

東京大学医科学研究所ヒトゲノム解析センター教授
理研ゲノム科学総合研究センター プロジェクトディレクター

さかき・よしゆき

1942年生まれ。1971年東京大学大学院理学系研究科博士課程修了。米国カリフォルニア大学ウイルス研究所研究員、三菱化成生命科学研究所副主任研究員、九州大学遺伝情報実験施設教授等を歴任。現在、東京大学医科学研究所ヒトゲノム解析センター教授、理化学研究所ゲノム科学総合研究センタープロジェクトディレクターを兼任。日本生化学会評議員、日本人類遺伝学会評議員、科学技術会議ゲノム科学委員会委員長。

国際プロジェクト「ヒトゲノム計画」に対して、スミソニアン賞「プラチナ技術21世紀バイオパートナーシップ賞1999」を受賞した。



<座長> 三輪 睿太郎
農業研究センター所長

親から子供に伝わるDNA

私たちは、親から子供にいろんな性質が伝わっていくことを経験的に知っています。例えばビクトリア女王の家系図を見ると、記録から血友病の性質が伝わったことを辿ることができます。親から子供に性質が伝わるといふことは、医学の面でもさまざまな文献的記載が蓄積しています。

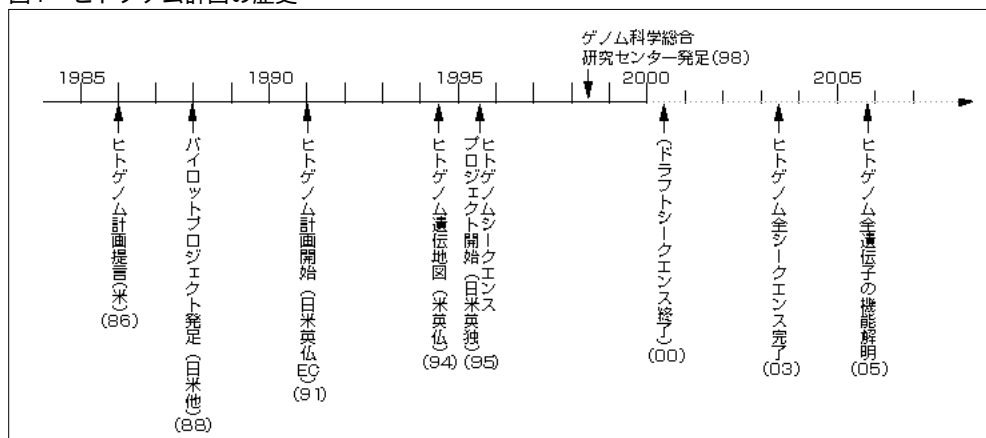
私たちの生命は、父親からの精子と母親からの卵子が一緒になることによってスタートします。この精子と卵子の中にはDNAという物質があり、それが親から子供への遺伝情報を伝えています。親から子に伝わる遺伝情報全体をゲノムといいます。それを伝えるゲノムDNAは染色体という単位に分かれています。人間の場合には1番から22番まで大きさの順に番号がつけられたものと、男女を分けるXとY染色体の合計24種類の染色体があります。我々はこの染色体を父親と母親から一本ずつもらうので、46本つまり23対の染色体をもっているわけです。染色体という情報をもらうことによって、親から子供にいろんな性質が伝わります。

染色体の中にはDNAという非常に長い分子がありま

す。DNAの中にはアデニン(A)、グアニン(G)、シトシン(C)、チミン(T)という4種類の塩基があり、AとT、GとCがペアを作って二重らせん構造をとりながら、つながっています。この中に親から受け継ぐ性質が入っているのです。ワトソンとクリックがDNAの構造を約50年前に決めて以来、どのような性質がどのように書かれているかを解くことが分子生物学者あるいは分子生命科学者の大きな目標でした。DNAはこれまでなかなか解読できなかったのですが、1970年代に開発されたDNA配列決定技術によって道が開かれてきました。

一卵性双生児は、どのような遺伝情報がDNAに書かれていて、どういった性質が環境に影響を受けているかということを知る上で、重要な研究対象になっています。例えば双子の「きんさん」と「ぎんさん」を見てもわかりますように、お二人とも非常に長寿で身長も体格も似ています。また性格という面でも、遺伝子が影響を与えていると想像できます。ですから、DNAによって私たちの体、あるいは性質の多くが決められていることがわかります。

図1 ヒトゲノム計画の歴史



ました。我々も当初これを発表すべきかどうかについて議論しましたが、ヒトゲノム解析を独自に進め、ビジネスとして利用しようとしているアメリカのCelera社が、解析の終了を新聞向けに一方的なキャンペーンを行っていたので、我々も解析経過を発表すべきである、ということになったのです。また、大雑把とはいえ遺伝子の9割方が

ヒトゲノム計画の歴史と進展

ヒトゲノム計画はDNAの中身を全部読み取ろうという計画です。1970年代に遺伝子解析技術が発達してヒトゲノムの研究が進んだことを受け、80年代にはヒトゲノム全体を理解することによって、人間を丸ごと理解できるのではないかという野心的なゴールが研究者の中に芽生えてきました。そして1991年から国際協力のもと、ヒトゲノム計画が公式に始まりました(図1)。

ヒトゲノムはA、T、G、Cという4種類の文字が、全体で約30億並んでいる文章のようなもので、その中に単語に相当する遺伝子という情報単位があり、そこにいろんな機能(意味)が含まれています。その遺伝子が相互に関係をもちながら働き、さまざまな生体反応を維持しながら、全体としては秩序だって働くプログラム(物語)が組み立てられています。ヒトゲノム計画はその30億の文字を決定して、どんな単語があり、それがどんな意味をもっているのか、どのようにつながっていて、全体としてどんなプログラムを組み立てているのか、を理解しようというものです。しかし30億という文字を手当たり次第読むのでは、どこに何が書いてあるのか全くわかりません。秩序だって読むために、まずゲノム上にいくつかの目印になる場所を探して、それを並べたものを作りました。それに基づき、順次30億の文字を読んでいく作業を1996年2月に開始しました。

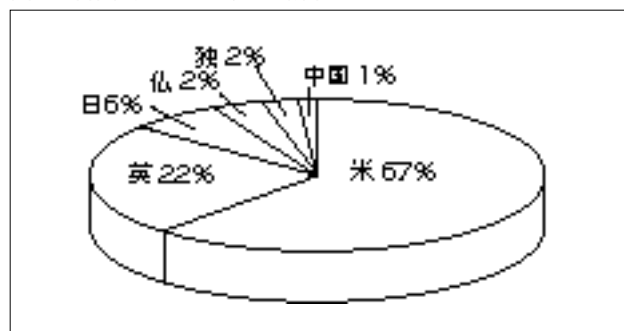
2000年の6月26日に、国際チームはヒトゲノムの9割を占めるドラフト、つまり全体を大雑把に読み取ったということを発表しました。サイエンスの立場からは、どうしてまだ概要の段階で発表するのかというご批判もあり

解析できたことは人間にとって非常に重要なことなので、皆さんに発表することになりました。アメリカではクリントン大統領、イギリスではブレア首相が出てきて大掛かりな発表をしました。しかし、強調したいのは今の段階ではあくまでも概要です。あと2年程かけて正確に解読しないと、サイエンスとしては価値が半減することです。

この中でデータ生産の割合は、アメリカが67%、イギリスが22%、日本が6%です(図2)。新聞では日本のデータ生産割合が「わずか6%」という報道がありますが、これは国の力の入れ方を反映した数字です。他の国と比べれば「よくぞ6%」とっていただきたい。今回の解析で完璧に読まれた染色体が二本あります。一番小さい21番と22番染色体です。21番染色体は、初めから日本が中心になって進めてきました。このプロジェクトの7割は日本で解析されたもので、日本が主導したものと国際的にも高い評価を受けています。日本の寄与は数字でいうと6%ですが、質的にはもっと高い寄与をしたと国際的にも認識されています。

しかし、日本にはもっと大きな貢献ができたチャンス

図2 各国別データ生産の割合



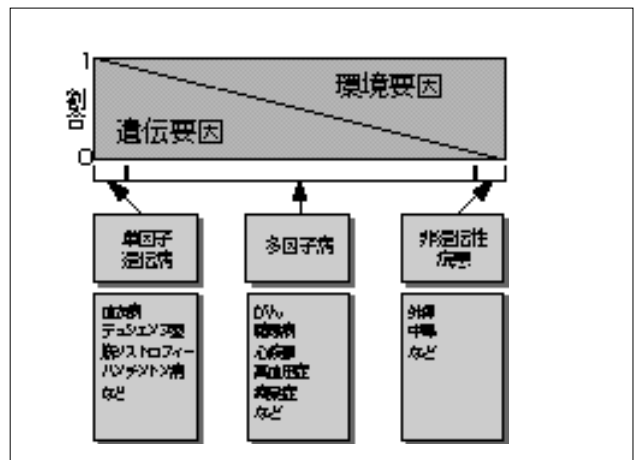
が2回あったと思います。初めは1987年、現在理化学研究所ゲノム科学総合研究センター長の和田昭充博士（当時東京大学教授）が、DNAの高速の解析システムを作ろうと日立やセイコー電子、富士フィルム等を集め、主要なパートについての各々の技術は相当レベルに達したということ、ネイチャーに発表したときです。これがアメリカを刺激しました。ワトソンはすぐにアメリカの議会に「このままだと遺伝子解析の分野でアメリカは負けてしまう。直ちに大量の予算を出して推進すべきである」と訴え、それまでアメリカの中でいろいろな議論があったヒトゲノム計画が、一気に推進するようになりました。さらに和田プロジェクトは完成をみないまま、途中で止まってしまいました。もう一つのチャンスは、遺伝子の解読をするDNAの自動シーケンサーという装置の技術開発のときです。この主要技術の一つは日立の研究所が持っていたのですが、日立がアメリカのPE社と技術提携したことにより、PE社が世界のシーケンサーの3分の2を支配してしまいました。この最終製品の開発は日立だけではできなかったかもしれませんが、日立がいなければできなかったでしょう。ただ、日本として独自に発展するチャンスを失った気がして残念です。

注目されるSNP研究

世界的な努力によってヒトゲノムが読まれ、30億の文字の中にいろいろな遺伝子があることがわかってきましたが、遺伝子と私たちの体との間にどんな関係があるかという研究は、まさにこれから始まります。ピクター・マキュージックというジョンスホプキンス大学の教授が、文献に書かれている人間の遺伝的な性質を次々と集めた辞書を1960年代から作っていました。これには現在1万2000件くらいの遺伝的な性質が入っていますが、遺伝子との対応がついているものはそのうちの1000件ほどです。残りの9割以上はまだ対応がついていないのですが、ヒトゲノムの解析が進んだことによってその対応が非常に早くわかってくるだろうと予想されます。

多くの病気や性質は環境要因と遺伝要因の重なりによって起こるので、単純に家系図を描いただけでは遺伝子との関係はわかりません（図3）。環境要因と遺伝要因の二つが重なった中で、遺伝的要因だけを取り出す新しい方法が最近有力になってきました。今までは家系の相関

図3



を遺伝学的に調べてきましたが、新しい方法は家系ではなく集団を扱います（図4）。ある病気の集団と病気でない集団を集め、遺伝子を調べます。病気が遺伝的な性質によって決まっているのであれば、病気の集団の方に特有の遺伝タイプが偏ることになります。例えばお酒を飲めるか飲めないかという違いは、アルデヒド脱水素酵素遺伝子のある一カ所が違うからなんです（図5）。アルツハイマー病になりやすいかどうか、アポEというタンパクのタイプによってわかります。このように我々の性質の違いは、DNAの塩基がたった一文字違うためにタンパク質のアミノ酸配列が変化し、それによってタンパク質の動きが変わることから生じています。このような違いを起こす一塩基の差（SNP；Single Nucleotide Polymorphism）が最近、注目を集めるようになってきました。

図4 遺伝子病の原因遺伝子の解析方法

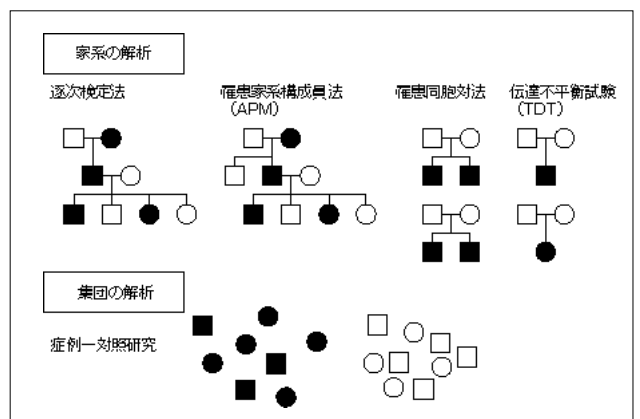
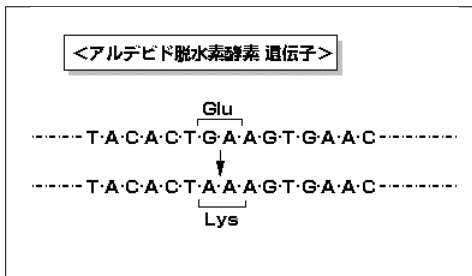


図5 アルコール感受性



ヒトゲノムの配列がほぼわかり、遺伝子が大方わかってきますと、SNPを大量に集めて

集団の相関解析にかければ、遺伝子と性質の関係が明らかにできます。現在ヒトゲノムの配列がほぼわかってきたので、次に個人差を示すSNPを探し出そうというプロジェクトが始まりました(表1)。日本はミレニアムプロジェクトの中で、5省庁が連携して日本人固有のSNPを見つけようというプロジェクトをスタートします。各国ともに昨年あたりから300万から500万といわれるSNP収集を行い、病気と遺伝子の関係を調べる研究に入っています。

しかし、予備知識なしにある集団と対象との間で遺伝子全体を調べるということは、非常に難しい問題です。そのため、ある性質と関わりの深そうな遺伝子について、SNPとその性質の相関関係を調べる研究が展開されています。例えば、血圧のコントロールに係る遺伝子はたくさん知られていますが、そのわずかな違いが高血圧になりやすいかどうかを決めていると思われるので、こういった性質とSNPとの相関関係の研究は特に活発に進んでいます。アンジオテンシノーゲンのある特定のアミノ酸置換は高血圧と関係があって、塩分に対する感受性が違うといわれていますが、この遺伝子タイプを調べることによって、塩分を控えれば高血圧にならないという人も出てくるのです。非常に意義深い研究ですが、今のところはまだ研究途上です。

遺伝子研究による個人情報の問題

SNPの相関解析のような研究には、対象と患者という二つの研究対象が必要ですから、最近では患者の合意を得るインフォームドコンセントが非常に大事になっています。日本の場合、遺伝子解析の研究が始まった当初に一部研究サンプルを黙って使用したことがあるため、ネガティブな印象を与えてしまいました。遺伝子の話は、サンプルを提供してくれた人たちに直接利益があるかな

いかという議論になってしまいますと、皆さんの合意を得ることは容易ではありません。個人情報きちんと守る、プライバシーを保護するということをしっかりと決める必要があります。

日本の中にも非常にいい研究対象はたくさんあります。例えば、九州の久山町には九州大学の第二内科が40年近くにわたって町民全体の健康調査をした記録とサンプルが残っています。それを使って研究をすればいい成果が出るのですが、今はそのサンプルに手をつけられない状況になっています。

現在、日本ではサンプルを集めるところに大きな困難があります。アイスランドではこういった問題を国として解決してしまいました。アイスランドにはバイキングが10世紀頃に渡ってきて、その子孫が代々住み続けました。現在の人口は約30万人ですが、祖先がわずかな数ですから遺伝的に均一な民族なんです。その中で特に高血圧、糖尿病になりやすいといった特別な性質があることがわかっています。国が全島民に遺伝情報の提供を求め、それをもとに病気のなりやすさを解析して、国の保健政策に反映させようとしています。もちろんたくさんの議論がありましたが、国会で法案を通して、現在国が作った会社を中心に遺伝子の解析を行っています。すでに遺伝学的に興味のある遺伝子がいくつか見つかってきています。アイスランドでは、遺伝子に関するさまざまな社会的問題に取り組んでいます。いずれ日本でも遺伝子に関する問題はますます広がってきますから、アイスランドでの試みは貴重な経験となってきます。

遺伝子の問題をさらに一歩踏み込んだものとして、生

表1 SNP情報収集に関わる動向

研究機関	概要・目標
GENSET社 (仏)	1999年までに150万個のSNPをマッピングする。60の異なる集団の代表となる30-40の多量遺伝子を見分ける。 ③年以内の中国を含む世界の健康調査と連携して、部分DNAと健康状態を把握。
The SNP Consortium Ltd. (東洋製薬と10社と東洋ゲノム研究所が共同で運営)	2001年までに300万個のSNPを単離し、その約15万個をマッピングする。 データは一部公開。 予算総額4,400万ドル(24億円)。
Celera Genomics社 (米)	③-5人分のヒトゲノムを解析し、3,000万個のSNPを抽出する予定。 ③500万ドル/5年でデータベース利用を認め企業に提供する予定。 ③100名のシーケンサーを約90人で24時間、③66日フル稼働させる計画。
CSNPプロジェクト (日本の5省庁が連携して遺伝子の協力研究)	CSNPを中心とする、③年間10万-15万個のSNPを特定、日本1500万個を探索、③年以内のSNP300万個を特定。

命保険会社の問題があります。遺伝子診断による病気の危険度の高い人は保険料を上げる、あるいは加入を拒否するといったこともできるのではないかという議論が世界的に起きています。日本でもつい最近シンポジウムが開かれましたが、非常に大きな問題になっています。イギリスの生命保険会社では、遺伝子診断を受けている人は報告して欲しいとっていますが、強制ではありません。アメリカはこういった問題を扱うこと自体、すでに規制をかけています。わが国では議論が始まったばかりですが、私自身としては絶対に規制すべきだと思っています。

遺伝子の解析によっていろんなことがわかってくると、医療やそれ以外のことに応用されるようになります。しかし、生命科学の場合には個人の究極のプライバシーといわれる遺伝情報に直接関わってきますので、議論を徹底的にした上で、次のステップを決めていくべきです。

遺伝子による薬物感受性と病気のメカニズム

ヒトゲノム計画が進んで、病気や遺伝的性質とその原因となる遺伝子との関係が付き、その情報を使った応用が可能になってきます。例えば、遺伝子診断の特色は、遺伝子に異常があるとわかれば、将来ある性質あるいは病気につながるということがあらかじめ予測できます。それが遺伝子診断のもっとも有用な点でもあり、もっとも危険な点でもあります。病気になる危険性を減らすことができる一方、治療対策や予防対策もないような病気の遺伝子をもっていることがわかって、本人にダメージを与えるだけです。遺伝子診断は強力な技術なので、どのようにうまく社会に適用させながら使えるか、ということが大きな問題になっています。診断する技術自身はそれほど難しいものではないので、問題は余計複雑です。

遺伝子診断が薬物代謝の遺伝子と薬物に関する感受性の関係を知るのに有効だということもわかってきています。チトクロームP450という遺伝子は酸化還元反応に関係しますが、特に抗がん剤の代謝と深く関わっています。遺伝子のタイプによって、抗がん剤がよく効く人、あまり効かない人、効きすぎて毒性が強い人に分けることができます。ですから、遺伝子のタイプを調べてから適切な抗がん剤を使う、といったことが臨床の場で行われ

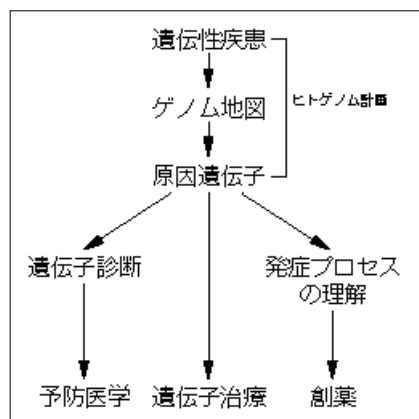
つあります。またNアセチルトランスという酵素は、結核の特効薬であるヒドラジドの代謝に関係しています。代謝が進む人には薬を使えますが、代謝が遅い人は肝臓の障害を起こすため、副作用が出てしまうことがあります。ですから遺伝子のタイプを調べることによって、薬を使えるかどうか判定できます。このように、薬に対する遺伝子のタイプによって、適切な薬を選ぶことが可能になります。

病気の原因である遺伝子がわかれば、直ちに病気が治せるというわけではないのですが、病気の発症メカニズムの解明が進むと期待されます。病気のメカニズムがわかると、どのステップを対象にどのような薬を開発すればいいか、どういう治療を施せば病気のコントロールができるかということもわかってきます(図6)。例えば、正常な粘膜が小さなポリープから大きなポリープになり、ガンになって転移していくという典型的な大腸ガンの進行には何万という遺伝子が関わってきます。これらの遺伝子がどのように発症のメカニズムに関与しているかがわかると、あるコントロールしやすいステップで発症を止める、あるいは転移を防げるといったことができるようになりますと期待されています。

ポストゲノムシーケンス：タンパク質の構造解析

最近、ポストゲノムシーケンスといった言葉が盛んに使われます。これはゲノムの配列がわかった上で、生体内で実際に起きているいろんな現象を遺伝子、タンパク質のレベルで理解しようという研究です。特にタンパク質の構造解析が、ポストゲノムの中でも非常に重要な

図6



ターゲットとして取り上げられています。今まさにヒトゲノムの配列を読み終わろうとしているわけですから、遺伝子同士、あるいはそれが作り出すタンパク質同士がどのようにして生体内の現象を作り出すか、ということが次の大きな課題です。国際的にもヒトゲノム計画に次ぐ国際プロジェクトとしてスタートしようとしています。日本でも「タンパク質の解析」にもっと力を入れる方向で動いています。

私たちが研究していた家族性アミロイドーシスという遺伝病の例ですが、この病気の原因物質であるトランスサイレチンというタンパク質のあるアミノ酸の種類が一つだけ変わると他の部分が物理的に押され、システインというアミノ酸が分子の表面に出てくることによって、この病気が起こりやすくなるのがわかっています。そのため、システインの動きを抑えることによって、病気の進行を遅らせることができるのではないかと考えられています。我々の研究の成果は、スウェーデンなどの臨床の場でテストされています。

生命のストーリーを再現するコンピュータ

遺伝子情報解読の上で重要なことは、1個1個の遺伝子がばらばらに働いているのではなく、相互に関係をもちながら秩序だって働いていることです。その働きが全体として我々の生命のストーリーを作り出していますが、それらはすべてA、T、C、Gの文字で書かれ、秩序だって起こることがわかっていますから、コンピュータ上に再現できると考えられます。コンピュータ上にゲノム情報をもとにした生体反応を載せ、シミュレーションで再現しながら生体反応を理解していこう、あるいは我々が今まで気がつかなかったような新しい規則性を見つけようというチャレンジングな学問が「システム生物学」です。この分野はまだ始まったばかりですが、アメリカではリー・フッド、日本ではソニーの北野宏明さんなどが提唱しています。

しかしごく小さいウイルスの遺伝子制御の系でも、コンピュータ上に再現するのは非常に複雑で困難です。ヒトゲノム全部をシミュレートできるようになるのは大変でしょうが、情報科学の発達によって将来は回路のようにゲノム情報が書かれ、それをもとにして生体反応をコンピュータ上でシミュレーションし、コントロールでき

る時代が来ると期待しています。

みんなで遺伝子問題の議論を

国としてはタンパク構造科学と同時にゲノム情報科学、生物情報科学をこれから大いに発展させるべきだという方向で動いています。特に、世界的にも人材が不足しているこの分野で、どのようにして人材を育成するかが問題になっています。教育改定ではDNAや遺伝子を教えないというような、時代に逆行する教育の「改悪」が行われるようですが、意欲をもった若い人が、自然にこういった分野に飛び込んでくることのできるような教育システムにして欲しいと思っています。

我々がゲノム研究を始めたときから、倫理的な問題をはじめさまざまな問題が出てくるだろうということは予測していました。我々は今までもワーキンググループを作って検討してきましたが、現実としているような問題が起きてきています。政府が生命倫理委員会を作って議論をしていますが、ゲノムや生命科学の研究者だけでなく、そろそろみんなで議論すべきレベルに来ていると思います。

理化学研究所のゲノム科学総合センターが2000年の11月6日に開所をしましたが、この時物理学者である前東大総長の有馬朗人先生が挨拶をされました。その挨拶の中で「物理学は20世紀の科学技術の時代を築き上げる中心になったと自負している一方、唯一胸を張れないのは原子爆弾を作って使ってしまったことで、まことに残念である。生命学者は、そういったことを二度と繰り返さないようお願いしたい」という話が出ました。遺伝子問題は、技術としては先鋭化していますが、その技術を使うことに関しては、非常にたくさんの議論を踏まえた上で進むべき問題だと思っています。

今日は21世紀の生命科学の素晴らしい展望だけを話せばよかったのですが、実はさまざまな問題を含んで出発しているということも、合わせてお話しさせていただきました。

SAT