

日本学術会議遺伝学研究連絡委員会報告  
—日本DNAデータベースの整備拡充の緊急性について—

平成5年2月25日

日本学術会議  
遺伝学研究連絡委員会

この報告は、第15期日本学術会議遺伝学研究連絡委員会で審議された結果を取りまとめ発表するものである。

第15期日本学術会議遺伝学研究連絡委員会委員

委員長	常脇恒一郎	(第4部会員・京都大学教授)
幹事	石和 貞男	(お茶の水女子大学教授)
幹事	福田 一郎	(東京女子大学教授)
	石川 辰夫	(帝京大学教授)
	大澤 省三	(名古屋大学名誉教授)
	五條堀 孝	(国立遺伝学研究所教授)
	坂口 文吾	(九州大学名誉教授)
	佐々木本道	(佐々木研究所所長)
	関口 睦夫	(九州大学教授)
	祖父尼俊雄	(国立衛生試験所部長)
	竹内 拓司	(東北大学教授)
	松永 英	(国立遺伝学研究所名誉教授)
	森脇 和郎	(国立遺伝学研究所教授)
	由良 隆	(京都大学教授)

## はじめに

DNAは、A、T、G、Cの4文字（個々の文字は塩基の種類を表す）で綴られた遺伝情報であり、すべての生物の形（構造）と働き（機能）を遺伝的に決めている青写真である。DNAが本体である遺伝子の全塩基配列を決定することにより、その働き、例えば、それがつくるタンパク質の構造を容易に推定できる。さらに、塩基配列の一部を改変することにより、遺伝子の働きを調節することができる。もし、一つの生物の全DNA（ゲノムという）の塩基配列を決定できれば、その生物が一生の間につくるタンパク質を全部把握できる。また、どの遺伝子が一生のどの時期に働くかを知ることにも可能となる。そのため、現在、多くの国において、いろいろな生物の重要な遺伝子の塩基配列が次々と決定されている。そればかりではなく、ヒトを含む2、3の重要な生物種については、その全DNAの塩基配列を決定するためのプロジェクトがすでにスタートしている。わが国においても、文部省や科学技術庁及び厚生省がヒトゲノムプロジェクト、農林水産省がイネゲノムプロジェクトに着手した。このような研究ないしプロジェクトが、小は研究者個人から、大は国家レベルにおいて盛んに行なわれるようになった理由は、DNAの塩基配列が、生物情報として極めて質が高く、その学問的価値はもちろんのこと、医療や多くの生物産業の将来にとっても、決定的に重要な意義をもつとみなされるからである。

ふつう、一つの遺伝子は、平均で数千の塩基よりなるが、それは、機能をもつエクソン部分と、働きを現す前に切り捨てられるイントロン部分とがモザイク状に配列したものである。また、各遺伝子に近接して、それが働くべき場所と時期を決めている、多くの特殊な塩基配列が存在する。その上、遺伝子と遺伝子の間には、現在の知識ではジャンク（がらくた）としか解しようのない長い塩基配列が見られる。ゲノムにみられる、このような多様な塩基配列から、意味のある配列を見つけ出す必要がある。また、世界の多数の研究室から時々刻々に生み出されてくる塩基配列を登録しチェックし比較して、重複を除き、新しいものを選別する必要がある。さらに、ある生物で決定された塩基配列の働きを知るには、他の生物で報告されている既知遺伝子の塩基配列と比較することが一つの有効な方法である。これはまた、生命の起源に関わる重要な知見を提供し、併せて、未知の遺伝子の存在や機能を予測することを可能にしている。

DNAデータベースは、このような必要性に応えるために考え出されたもので、塩基配列とその周辺情報（遺伝子の名前とその定義、遺伝子の各部位の説明、その生物学的機能の説明、研究者と掲載論文の名など）をコンピュータ化したものである。個々の研究者が、自分のパソコン等から電話回線やコンピュータネットワークを利用してこのDNAデータベースにアクセスすることにより、自分の解析した塩基配列の新規性や既知遺伝子との比較、あるいは知りたい遺伝子についての諸情報の入手が、即刻可能である。また、これを利用して、自分が関心を持っている問題を研究している研究者に接触することも併せて可能である。

このようなDNAデータベースの利用価値は、その収蔵している情報量とその同時性（鮮度）に大きく依存する。そこで、すべての生物の、すべての遺伝子について、世界のすべての国で報告された塩基配列を、即刻データベース化することが要求される。このような役割を担っているのが、米国のGenBank、欧州のEMBL Data Libraryとわが国の「日本DNAデータバンク（略称DDBJ）」の三機関からなる「国際共同DNAデータベース構築機構」（図1）である。

日本学術会議遺伝学研究連絡委員会においては、国際的なDNAデータバンク事業の現状について検討を行ない、わが国におけるDNAデータバンクの問題点と将来の課題を明らかにしたので、ここにその検討結果をとりまとめて報告する。

## 1. DNAデータバンク設立の経過と国際的状況

わが国においては、DNAデータバンク研究事業を、国立遺伝学研究所の遺伝情報研究センター遺伝情報分析研究室が、1984年より、「日本DNAデータバンク（DDBJ：DNA Data Bank of Japan）」の名称のもとに担当することになった。

欧州では、これより早く1980年に、EMBL（欧州分子生物学研究所：European Molecular Biology Laboratory）がEMBL・DNAデータバンク（以下EMBLと略称）を設立した。次いで、1981年に、米国で、国立ロスアラモス

研究所が中心となって、「GenBank」という名のもとに、DNAデータバンク研究事業を開始した。さらに、1992年になると、NIH（National Institutes of Health）のNCBI（National Center for Biotechnology Information）が「GenBank」を担当することになり、その活動を強化した。

「まえがき」に述べたごとく、DDBJ、GenBank、EMBLは、全世界的に、DNA配列データの収集・管理・提供を行なうため、「国際共同DNAデータベース構築機構」（図1）を構成し、国際的に業務を分担している。特に、DDBJは、主に日本国内におけるDNA配列データの収集を行ない、各種データベースの提供を分担している。これら三者は、国際協調を円滑に行なうため、DNAデータバンクの国際諮問委員会と国際実務者協議会を設け、それぞれの会合を、年1回、日・米・欧の持ち回りで開催している。

このように、日本DNAデータバンク（以下、DDBJ）は、GenBank、EMBLと協同して、DNA塩基配列データの収集・管理を行ない、両DNAデータバンクとデータ交換を公式に行なっているわが国唯一のDNAデータバンクである。

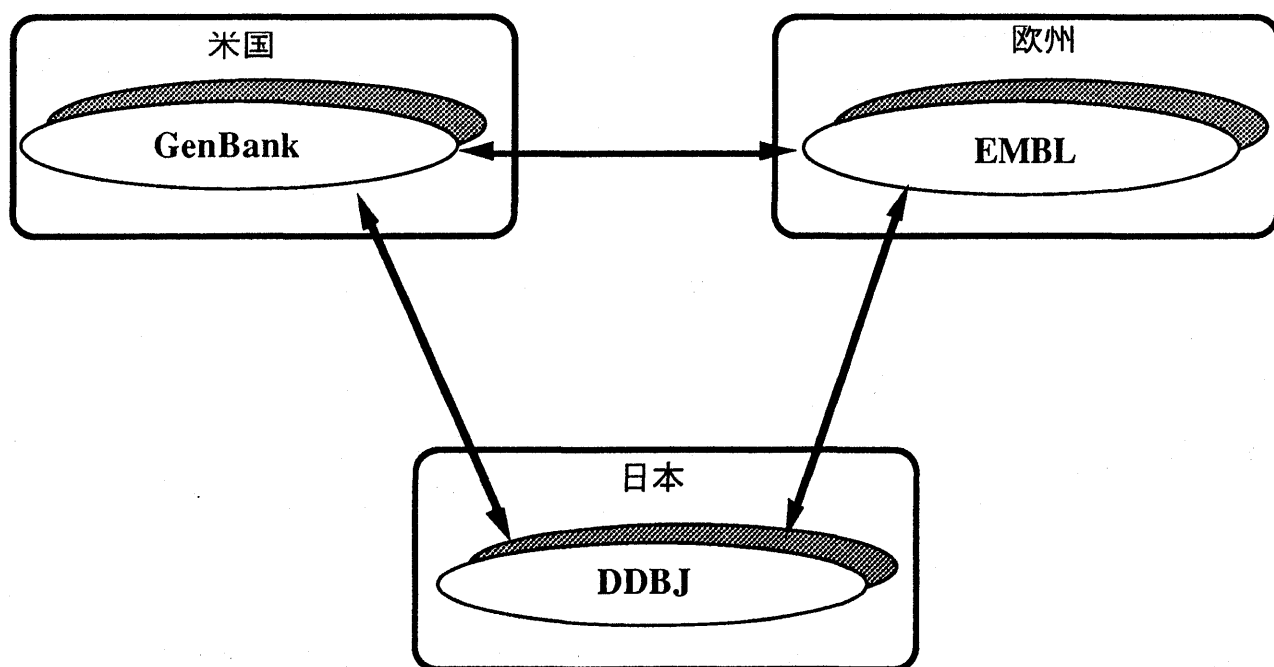


図1. 国際共同DNAデータベース構築機構

## 2. 国際的 3 D N A データバンクの事業規模の比較

DDBJ、GenBank、EMBLの国際的 3 D N A データバンクの、1992 年 9 月現在の年間の予算と入力件数、及び組織の規模を表 1 に示す。DDBJの年間予算は、事業費、スタッフの給与、及び共用コンピュータのレンタル料を合わせた金額である。

表 1. DDBJ、GenBank、EMBLの事業規模の比較

1992年9月現在

区 分	DDBJ	G e n B a n k	E M B L
予 算	約1億5千万円	約54億円	約35億円
入力件数	約1,200件/年	約7,000件/年	約7,000件/年
組 織 (正式スタッフ)	7人	約50人	約35人
担当政府機関	文部省	N I H	E C 政府

表 1 にみるように、米国及び欧州の D N A データベースに比べ、DDBJは、予算規模において 1/23 - 1/36、入力件数において 1/6、正式スタッフ数において 1/5 - 1/7 に過ぎない。DDBJは、その予算規模及びスタッフ数からみれば、きわめて効率良く事業を実施しているといえる。この健闘は大いに讃えられねばならない。

しかし、国際協力の観点から眺めるとき、日本の貢献（データ入力数）は全体の 8 % 弱に過ぎず、わが国研究者の国際共同 D N A データベースの利用頻度に比べ、極めて低い。そのため、関係者の懸命な努力にも拘らず、欧米から日本の「データただ取り論」が浮上してくる危険性が日に日に高まっている。

### 3. 日本DNAデータバンク (DDBJ) の整備・拡充の緊急性

DNA配列データが、科学的価値ばかりでなく、その経済的価値も次第に強く認識されるようになったため、国際共同DNAデータベース構築機構に対する、わがDDBJの均衡を失した、低い貢献度は、単なる「データただ取り論」を越えて、経済摩擦の様相を帯びる恐れがある。その結果、DDBJが、国際共同DNAデータベース構築機構から排除され、わが国の研究者が、現在のように自由に、国際共同DNAデータベースを利用できなくなる事態が予測される。これは、わが国のDNA関係の研究者のみならず、生物産業全般にとっても致命傷となるものであり、このような事態の招来を避けるため、緊急にDDBJの整備・拡充を行なう必要がある。

一方、国内における利用状況に目をやると、DDBJのデータ入力数は、図2に示すように、増加の一途を辿っており、過去1年間の平均で、入力数は毎月約100件になってきている。これに応じ、GenBank及びEMBLへのデータ送付件数も着実に増加している。また、DDBJを利用するオンライン登録者数も、図3のように急増しており、1,000の大台を越えるのも時間の問題である。

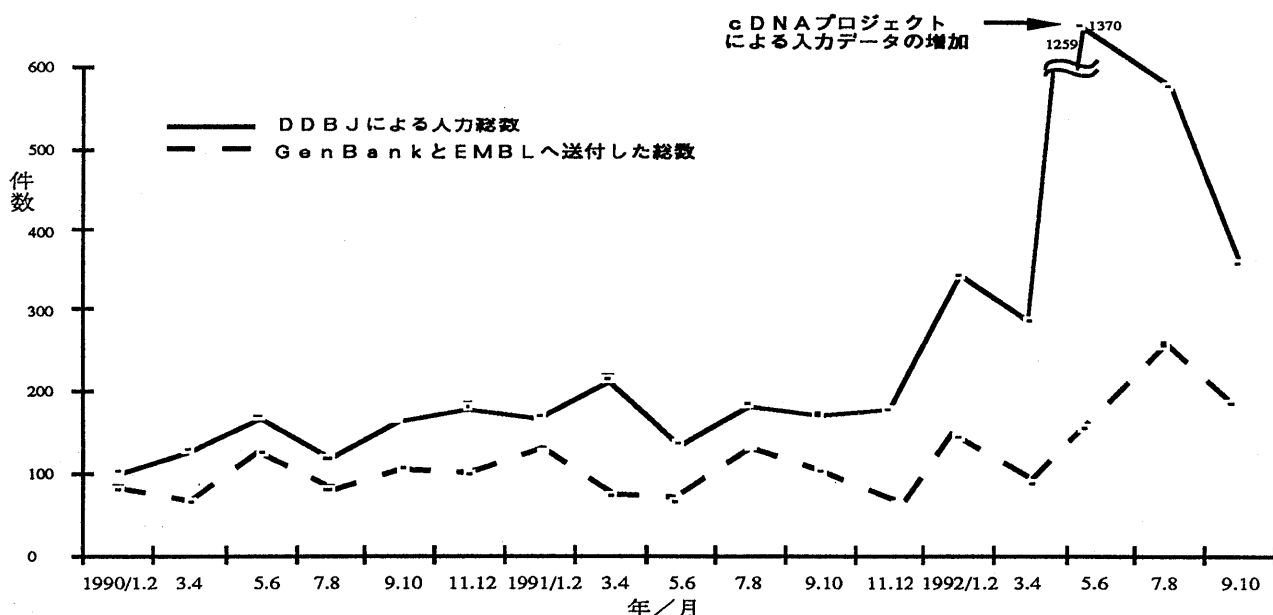


図2. DDBJによるデータ入力件数とGenBank及びEMBLへのデータ送付数の推移

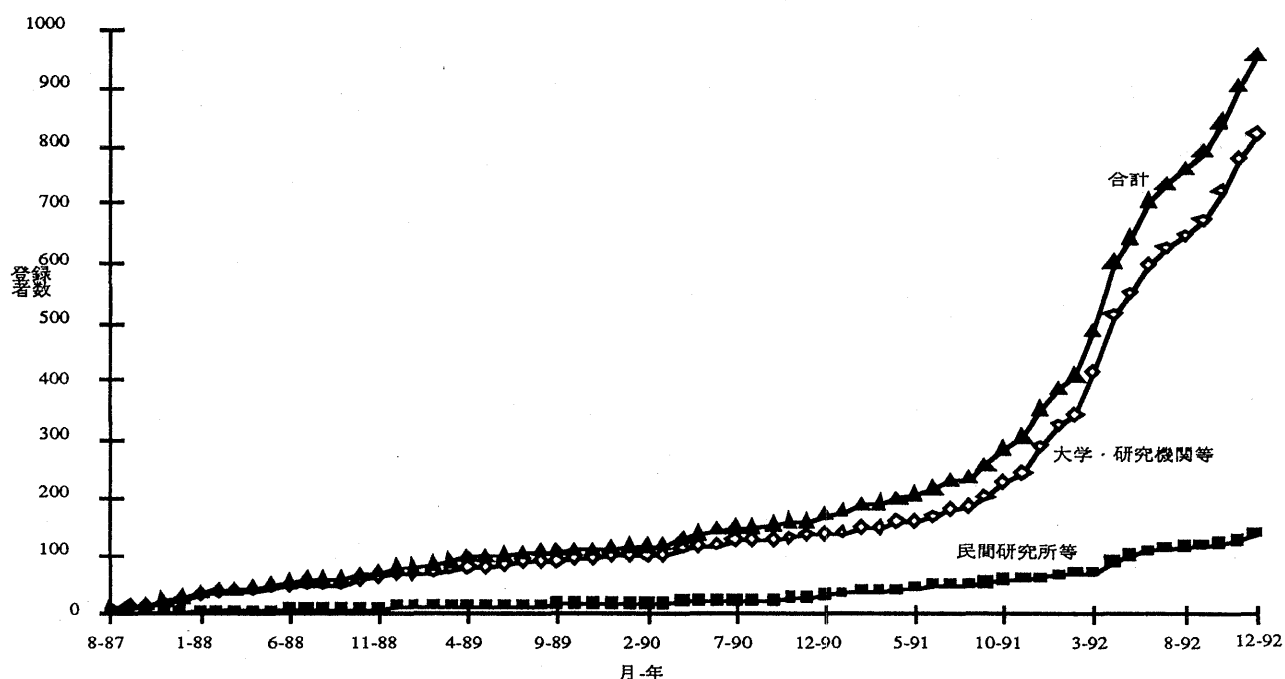


図3. DDBJのオンライン登録者数の推移

さらにまた、最近スタートした二つのナショナル・プロジェクト、「ヒトゲノムプロジェクト」と「イネゲノムプロジェクト」の進展により、今後、DNA配列の情報量が桁違いのスピードで増加することが予想される。このような国内の急増する需要・要請に対して、DDBJがその任務であるDNA配列の収集・管理・提供を迅速かつ円滑に行なうためにも、その機構と予算を重点的に拡充・強化する必要がある。

このような、国の内外からする要請にDDBJが必要かつ十分に対応できるようにするには、当面の措置として、年間3億円の研究事業費と年間6億円のコンピュータレンタル料の確保、及びスタッフ10名の増員を、今後2年内に実現する必要があると判断される。また、人材育成の一環として、DDBJのためにポストドクトラルフェローを常時5名程度確保するための措置を講ずる必要がある。

#### 4. 日本DNAデータバンク (DDBJ) に関わる今後の課題

現在、わが国の大学や研究所等では、生命科学と情報科学の教育が別個に行なわれており、これら両分野やその境界領域を体系的に教育する機関が殆



どない。このため、DNAデータバンクに必要な研究事業を担える人材の育成が困難な状況にある。従って、大学や研究所に生命情報科学の講座・部門等を新設する必要がある。

また、DNAデータベース構築には、生命科学についての高度な専門知識と、急速に発展する情報科学の思考法と方法論を駆使する、先端的な開発研究が必要である。一方、高度に特殊化したDNAデータベースを一般の研究者が活用するには、そのソフトウェアの公開や、データベース利用のマニュアルの提示など、きめ細かい研究支援サービスが必要である。このような研究開発とサービスを効果的に行なうには、一つの研究センターにこれらの活動を集中させることが望ましい。

そのための具体策として、DDBJを中核にして、これらの高度な開発研究とデータサービスを一体的に行なう「生命情報研究センター」を、できるだけ早い機会に、全国共同利用研究機関である国立遺伝学研究所内に設置することが適切と判断される。

## むすび

上述したように、急増する国内外からの日本DNAデータバンクに対する需要・要望に適切に対処し、わが国に課せられた国際的役割を応分に果たし、かつ、国内の生命科学と生物産業の発展を支援するには、日本DNAデータバンクの早急な整備・拡充が必要である。