

DNA Data Bank of Japan

国立遺伝学研究所

遺伝情報研究センター

遺伝情報分析研究室内

目 次

巻頭言	3
EMBL/NIHワークショップ「分子生物学におけるデータベースの将来」開催される	4
DNAデータ入力：DDBJ Release 1.0	14
Nucleic Acid Research の投稿者へ：EMBLへのDNAデータの提出に関して	21
現在データバンク間で協議されている事項	24
DDBJ 計算機システムのオンライン利用開始	26
DDBJ に収集されている配布可能なDNAおよび蛋白質データベース	27
DNAデータベースの配布に関する活動報告	33
データバンク利用者へのお願ひ	36
DDBJアンケート	37
DDBJ関連行事日程表	51
編集後記	52
訃報	53
弔辞	54

DDBJ計算機利用の手引

付属資料

DDBJ/GenBank/EMBL Data Submission Form

DDBJニュースレター申し込み書

DNA,蛋白質データ配布申し込み書

ソフトウェア配布申し込み書

国立遺伝学研究所DNAデータベース等利用申請書

国立遺伝学研究所DNAデータベース等利用終了、中止、承認内容変更届

Mailing Addresses for Inquiries

If you have any inquiry, please send mails to the following addresses (...@niguts.nig.junet).

nig	about the nig system
postmaster	about mails, including address representation.
ddbj	about DDBJ activities
ddbjsub	data submission to DDBJ
genbank	inquiries to GenBank
gbsub	data submission to GenBank
embl	inquiries to EMBL
emblsub	data submission to EMBL

Mails to genbank, gbsub, embl, and emblsub will be forwarded to them.

NOTE: On the nigvms, type "\$ smtp ...@niguts.nig.junet".

DNA Data Bank of Japan 所在地

411 三島市谷田1111
国立遺伝学研究所
遺伝情報研究センター
遺伝情報分析研究室内 DDBJ
電話 0559-75-0771 (内) 647
E-mail: ddbj@niguts.nig.junet

DNA Data Bank of Japan スタッフ

General Manager	遺伝情報研究センター長	丸山毅夫 (tmaruyam)
Manager/Database administrator	遺伝情報分析研究室	宮沢三造 (smiyazaw)
Scientific Reviewer	遺伝情報分析研究室	林田秀宜 (hhayashi)
関連事務担当	DDBJ 非常勤	堀江元乃 (dbs)

(E-mail:)
(@niguts.nig.junet)

巻頭言

遺伝情報研究センター長 丸山毅夫

DDBJ (DNA Data Bank of Japan)では去る7月、自力で構築したDNAデータベースの第一版を出しました。これは主として日本で出版されたデータを編集収録したもので、欧米 (GenBank、EMBL) との国際協力の一環をもなすものです。まだ GenBank (米国) や EMBL (欧州) に比べて小さい規模ですが、DNA データベースは現在の分子生物学の研究には不可欠なものであり、日本独自の構築は大変意味のあることだと思います。これは DDBJ としてアクセスでき、また、DDBJ 1.0 として利用者へ配布することも出来ます。また DDBJ 1.0 は GenBank と EMBL にも送られていますので、それらのデータベースの新しい版には入ることになります。

DDBJ へのオンラインアクセスが可能になりました。通常の話回線とパケット回線 (DDX) の両方法が可能ですのでご利用下さい。なお、これを機会に「DNA データサービス案内」を作製し希望者と関係機関・学会などに配布し、データベースの利用の普及に努力しています。

今年2月25-27日ハイデルベルグのEMBLでEMBL/NIH主催のワークショップ「分子生物学におけるデータベースの将来」が開かれました。近い将来に予想されるDNAデータベースの爆発的増加と利用形態の劇的な変化に如何に対処すべきかが、コンピュータ、情報科学、分子生物学、行政までの幅広い分野の専門家によって、いろいろな面から意見の交換と熱心な討議が行われました。その際、日本において最近関係官庁がDNAデータベースに強い感心を示し、国際協力の準備が整って来ていることが大いに歓迎されると共に日本からの寄与が日本におけるライフサイエンス分野の能力に見あう規模でなされることが要望されました。なお、ワークショップの報告書の一部抜粋を本号に載せています。

DNAデータベースを今後も順調に維持するためには著者の協力が欠かせない条件となる事は上で触れた国際ワークショップでも再三強調されました。DDBJでも昨年からのデータ入力に関して著者の協力を得たい旨、関連分野の学術雑誌編集部へ依頼していましたが、そのすべての編集部から返事を頂き、編集部としての協力をしてくださる事になりました。著者の皆様へは編集部を通して依頼が行きますので、その趣旨をご理解いただき、是非ご協力下さるようお願いいたします。

人事の面では、昭和61年度遺伝情報分析研究室に助手一名のポストが認められ、公募による人選の結果、林田秀宜 (九州大学・院) が決定し、4月1日付赴任しました。

丸山 毅夫

今年2月25-27日、西独ハイデルベルグのEMBLにおいて標記のワークショップがEMBLとNIHの主催で開かれました。ワークショップは、近い将来に予想されるデータベースの爆発的拡張と利用形態の劇的な変化に動機付けられたもので、参加者は約50名、参加者のスペクトラムはコンピュータや情報科学の専門家から分子生物学の実験科学者や政府の行政官に及ぶ幅広く、協議・討論も極めて多岐にわたりました。参加国は米国をはじめ欧州のEMBL関係国と日本で、わが国からは内田久雄（帝京大）・金久実（京大化研）教授と私の3人が参加しました。ワークショップのあと出された Press Release、ワークショップの要約、勧告及び一般問題についての討論の摘要を下に載せます。なお、ワークショップでは、このデータベース構築の事業に日本が積極的に参加すべく科学者と関係官庁が準備していることを歓迎すると共にそれに大きな期待が寄せられました。今後DNAデータベースの重要性が益々増加することはほぼ自明なことです。その運営が円滑に運ぶか否かは国際協力に大きく関わっているように思われます。その様な事情に鑑み米国・欧州・日本からなる国際アドバイザリー委員会の設置が会議で勧告されました。

EMBL/NIHワークショップ「分子生物学におけるデータベースの将来」、EMBL
ハイデルベルグ、1987年2月25-27日、報告書の抜粋

EMBL Press Release

1987年 3月 2日

”次世代（5年後を想定）における生物学データベース”と題する国際会議が1987年2月25-28日、ハイデルベルグのEMBLで開催された。参加者は欧州、米国、日本から分子生物学者とコンピュータ科学に従事するものが集まった。本会議は、EMBL (European Molecular Biology Laboratory) および、NIH (米国National Institutes of Health) の国立一般医科学研究所 (National Institute of General Medical Sciences) によってサポートされた。データベースの将来像と必要条件に関する今回の会議はデータベース使用における劇的な変化と予測されるデータベースの爆発的拡張という2つの事実の動機づけられたものといえる。現在のデータベース、特にEMBLデータライブラリーとGenBankが維持している核酸シーケンス・データベースはもうすでに、こうした急速な変化によってもたらされる幾つかの問題に直面している。ヒトの全ゲノムのシーケンスは発展しつつある技術の手のとどくところに来ており、最近ではこのような大きな規模で

生産される、関連した情報とシーケンスを処理し維持する方法への必要性がいたるところで強調されている。将来の生物学研究において必要とされるデータ保管、データ転送、データ処理技術の発展について討議された。たとえば、通信技術の発展は、生物学データベースの将来にとって不可欠な条件である研究者自身によるデータの維持と提供という共同責任を可能にするであろう。5年間に及ぶ、EMBLとGenBankの協力体制の成功は、更に拡大した国際協力への可能性を示す重要なモデルケースといえよう。これを支持する意味で、国際諮問委員会の設立が勧告された。委員会は分子生物学及び情報科学の専門家からなり、データベースの発展において現在の重要な時期に助言と指針を与える。データベースの前途には脅威的な仕事がかかっているにもかかわらず、科学者社会からのサポートと必要な技術知識の応用によって生物学データベースの将来は確かなものになりうると思われる。

1 歴史(省略)

2 会議の形式(省略)

3 要約(summary)

要約(summary)のセッションで次の事柄が重要な事として明確になった。

データベースの将来、特に今後5年間を予測する上で実際の問題はシーケンスデータ量の予測しがたい率での爆発的な増加である。この爆発的増加は我々のデータベースに関する概念と必要度を大きく変えるであろう。

データベースが、発表されるデータの入力に追い付いて行くためには、研究者がシーケンスに注釈をつけ、コンピュータ可読な形で直接データベースへ提供することが必要であろう。このような事が滞りなく行われるための学術雑誌編集部からの協力も必要となるであろう。最近データベースと学術雑誌編集部との間で取り交わされている協力関係はこの点で大変注目値する。このような協力関係への努力は今後一層広げられるべきである。さもないければ、データベースは最小限の注釈をつけ、研究者の要求に対して最小にして最も重要なものだけを提供するという事を続けなければならないであろう。

シーケンスデータの量が増加するにつれ、学術雑誌の編集部はシーケンスデータの論文を発表することに益々躊躇するようになるであろう。にもかかわらず、それらのシーケンスデータは入手可能な形の情報としてデータバンクに入れるべき重要なものである。こうした状況にあって、シーケンスデータの投稿を奨励促進するために、また科学の世界に必要な学術誌における研究成果の発表という信用機関として新しいシーケンスデータを載せる”学術雑誌”が必要かも知れない。

1992年には、分子生物学者がコンピュータの恩恵に預かる面が今よりずっと大きくなるであろう。このような発展にはいわゆる使い易いソフトウェアが要求される。将来の科学者は、科学者間のコミュニケーションのため、またデータのコミュニケーションのためコンピュータネットワークを使うことが出来るようになるであろう。これらのネットワークはデータベースにとってもいろいろな面で役立つ。たとえば、利用者からネットワークを通してデータベースのマネージャーへフィードバックすることなどが可能になる。ニュースレター等の方法によってデータベースとの短かな通信は増加することができる。

科学の進展につれ、ときどきデータベースに新しい事項、すなわち分子生物学が関連した他のデータベースとの連結や、最も重要なのはデータベースマネージャーと利用者の必要性に応じてデータベースの構造に新しいものを採り入れていく事などが必要となろう。

上述したような必要事項を満たしていくためには、どうしても緊密な協力が継続されることが基本となるであろう。現在行われているGenBankとEMBLデータベース従事者との間の協力は、この種の協力が実際に可能であることを示すモデルケースであり、将来の国際協力が、この事に関心を寄せるすべての者にとって如何に有益なものであるかを示す理想像となるべきであろう。日本のデータベースプロジェクトの加入は大いに奨励されるべきである。しかしこのプロジェクトは現存するものと密に調整されていかなければならない。

4 勧告

本集会参加者は、(分子生物学)データベースが将来起こると予想される急速かつ予測困難な変化を吸収し、変革することを援助、実現するために、データベース活動について各々の担当機関に助言する国際的委員会を設立することを強く要望する。この委員会は米国より3名、欧州より3名、日本より2名の委員から構成されると予想される。この委員会はまず核酸塩基配列データベース(の発展)について面倒をみる。この委員会は効果的であることが必要とされる科学的専門知識の幅と深さの両者を持たなければならない。委員は米国においては学士院、欧州においてはEMBO、日本においては類似の組織の推薦によって指名されることが期待される。この委員会は将来の問題に対処するための変革が協調をもって解決されることを保証するであろう。

このパネルは幾つかの幅広い件案について討議する：

- 一 増加の一途をたどっている必要性（データベースの）の性質と規模はどうであるか？
- 一 必要とされる努力（プロジェクト・業務）に対して誰が（どこが）資金を出すのか？
- 一 如何なる動機づけ（この事業に対する）によって必要な政治的支持を獲得できるか？
- 一 努力がどのようにして組織されるか？

チャールス デリーシ (C. Dilisi) 氏はこれらの問題の第一番について特に重点をおいて話した。(彼の”データベースの需要についての予測”と題する別紙の報告を参照) シークエンスの速度が10,000,000ベース/日というような速度に増加するとデータ処理に質的な変革が要求されるであろう。つまり、シークエンスの特徴づけ(この場合、特定の機能や役割をするベースの同定など)やアノテーション(注解)作成は自動化されなければならない。直接著者によってデータを入力し、注解作成を完成させることにより、この問題は部分的に解決されるが、しかし何らかの自動化が行われなければ、データの入力と編集に要する費用がシークエンスすること自体に要する費用に比べて不釣り合いに高くなるであろう。後者(シークエンシング)の費用は1990年までにはおそらく0.17ドル(US)/ベースにまで低下するにちがいない。

シークエンシング速度における、このようなかなりの加速度的増加がもたらす知的挑戦、ヒトの全ゲノムのシークエンシングのビジョン、それによってヒトという”種”の基本を決定し、ヒト個人の独自性を決めるような研究は、必要な政治的支持を得るだけの一般市民の関心をひきつけるに十分なものかもしれない。このようにして作られるデータベースは大変貴重なものであり、恒久的資源といえよう。より世俗的な目標としてあげるなら、人間の発達(発育)過程の解明や病気への貢献、バイオテクノロジーにおけるよい意味での競争の基礎づくりへの貢献などがあげられる。

コスト(シークエンシングの)は高いようにみえるかもしれない(0.17ドル/ベースの単価で10,000,000ベース/日と計算すると、5000万ドル/月)。しかしこれは遺伝病を科学の力で治すことによって、増加の一途をたどる病気の医療費の一部を社会から軽減することを考えると、むしろひかえめかもしれない。たとえば米国には70000人の糖尿病患者があり、その半数は若年者である。これらの治療に要する費用は年間10億ドルにのぼる。もしヒトのゲノムプロジェクトが若年者糖尿病発病の引金である抗原と自律抗体反応について、その鍵を解明するようなことがあれば、このプロジェクトはここにあげた病

例だけでも正当化されることになる。このような例は他にたくさんある。

くり返し強調された点は長期にわたる安定した資金 (funding) のサポートを確保することの必要性であった。 データベースへは今までのところ、生命科学をサポートしている研究費のほんの僅かな一部分しか回されていなかった事が指摘された。

データベースの重要性 (役割の) の認識が科学者の中で質的に変わった。この認識の変革は研究費を担当する機関 (funding agencies) にも反映されてきた。また、データベースは科学の手法をも変えた。いまでは、科学者にとって、それが学術分野か産業分野かを問わず、効率の良い研究のためには、データベースは基本的に必要なものとなった。したがって、担当機関 (funding agencies) は狭い意味での”研究”のみをサポートするのではなく、データベース、協力活動、ネットワークの基盤となるような研究をサポートすることも必要である。トレーニングの面でも同様なことが言える。 データベースにみられる質的变化は、このワークショップにおける他の発表にも現れている。たとえば、Ma sys氏の報告 (paper) は分子生物学データベースと学術雑誌の関係の変遷について述べている。

配布されるデータベースシステムが必要とする組織について研究することが必要である。これに関し必要な国際協力を促進するための技術的開発が必要である。データベース自体、コミュニケーションの重要なチャンネルの一つになりつつある。

データベースは学術雑誌の単なる電子版としての役割をし、科学活動全体の中の普通の単なる一部分と見なされるべきか、または、データベースのサービスの性質がこのようなものから意図的に少し離れるに値するもので意見が分かれた面があった。EMBLにとっては産業界や他の共同出資金を得ることが比較的容易であり、また、欧州バイオ情報サービス (European Bio-Information Service) を通じて核酸シーケンスを配布することによる収入は容易であろうとの助言があった。他方では、シーケンスデータベースは関係した研究と物理的に近い関係を保つことが強調された。勿論、このような関係をシーケンスデータに潜在的に関係のある全ての生命科学の分野と保つことは不可能である。しかし、シーケンスデータベースの発展自体かなりの研究活動を含む。

このような考えの精神に沿って、Philipson氏はEMBLにおけるシーケンスデータベースと生物科学計算の理論研究 (C. Sander氏指導) の展開について、その概略を述べた。

データベースと利用者の関係についての提案として、F. Ruddle氏のヒトの遺伝子地図ライブラリ (Human Gene Library) の状況と利用科学者が発展と新企画について討議する年間2回の会議 (conference) が参考にあげられた。しかし、この場合も、他の多くのデータベースの場合も、その目標や目的は狭く、利用者集団もシーケンスデータベースよりずっと狭いものになることは最後のセッションで更に討議された。

このワークショップの国際的性格は大変有意義なものであった。米国、日本、欧州から相当な参加があった。将来のシーケンス努力の規模とデータ、ソフト、アイデアの交換の必要性をみると国際協力は欠くことのできない基本的なものである。如何にして、この多岐多様にわたるfundingを組織したらよいのだろうか？ この問題については多くの議論が交わされた。状況は米国、日本、欧州と国によって大きく違っていることは明白である。データベースのこの点に関する要素（要点）は付録として後記する。この討論から出てきたことを次ぎのように要約できる。

米国では多くのデータベースが開発されている。それと共にいろいろな機関が資金(funding)を増加するよう努力している。国内的に国際的に利用者が幾つものこうしたデータベースにアクセスできるためのゲートウェイの開発にNLMが果している役割は、利用者が直接行っているデータベースへのアクセスに加えて貴重なものと言える。

日本については言えば、日本でデータベースに関する関心が高まり、幾つかの省庁がそのサポートに興味を示していることをこのワークショップは注目し、このことを大変歓迎する。その重要性を強調し、米国と欧州の参加者は次ぎのように強調した。データ交換と他の面での協力を調整し促進する日本のセンター（ズ）が発展することを期待する。それは特に日本で生産(generate)されるデータに関する国際共同体への協力（国際協力）を司り、生命科学における日本の強力な、そして成長しつつある能力に応じたものであることを期待する。

欧州では、EMBLの中心的役割は、西ヨーロッパにおける分子生物学データベースの自然な焦点として認知された。その生物学計算の役割とサービスセンターとしての役割はいろんなレベル（地域的、国家的、EEC）の政治的当局のサポートを得て拡張されるべきものである。安定した長期fundingの確立がデータベースの信用と権威を実証するであろう。利用者、たとえば産業界、からの寄付などについてもよく調査してみる必要がある。

EMBLや国のセンターにおけるデータベースの組織、入力と配布、ソフトの開発などのそれぞれの役割は、技術の変化と共に変わってゆくであろう。しかし、いつもその役割は明記されていなければならない。これは利用者だけでなく、政治的な意味でも言える。

米国、日本、欧州における分子生物学のデータベースへのfundingと組織

アメリカ

米国については、それぞれの省庁の参加者がそれぞれの方針について話した。エネルギー省(Dilisi氏)は大規模で長期にわたるプロジェクトの運営の経験がある。NIH(Cassatt氏)は複数の省庁出資によるGenBank (NSF, DOE, NLF, USDA)について述べた。Ge

nBankは政府が作ろうと欲したもので、契約(contract)によって行っているものである。従って研究費(grant)とは異なる。研究費の場合は普通、研究者がイニシャチブを採る。Gen Bankは、タンパク同定リソース(protein Identification Resource)と同様に、研究資源のNIH部門によってサポートされている。C. Coulter氏によれば、上記研究資源は全体として年間300-350万ドルを直接データベースに、そして250万ドルをその関連事業に出している。GenBankは今までに、350万ドルかかっており、その間のNIHの全体予算は250-300億ドルであった。タンパク構造に関する費用は年間2000万ドル、DNA構造のそれは1600万ドル、結晶学は年間2000万ドルであった。したがって、今までそれがサポートした研究の内容からみて、GenBankはバケツの中へほんの一滴といった感じである。

J. Wooley氏は、NSFは他の省庁よりmission oriented 型ではないことを述べた。しかし、NSFはデータベースに対する認識を変えた。この方面へのサポートを強化することを決めた。生物科学におけるNSFの増大しつつある役割のもう一つは、スーパーコンピュータセンターと生物学機器開発センターのサポートである。生物のデータベースのサポートに関するアメリカ議会への要求は系統センターと関連づけられるであろう。

G. Cahill氏は、ハワードヒュー医学研究所(HHMI)は”財団foundation”ではなく”組織organization”であり、病院と協力して研究を行っていると述べた。IRSはリソース(資源)を支持するためのサポートを拡張することを認めた。分子生物学の分野では、それらのサポートは高度な臨床遺伝学から基礎的な核酸の研究まで含まれる。ニューヘブン(New Haven)では、ヒトの遺伝子地図を国(実際には国際的)のリソースとしてサポートしている。ジョーンズホップキンス大学(Johns Hopkins)には、McKusickのヒトのメンデル遺伝データベースがある。これらのデータベース間の連鎖結合を促進し、またマウスの遺伝子地図データベースとの結合も望まれる。これらのデータベースを編集する小さなノード(node)はサポートされている。

パリの多型(polymorphism)研究センターもサポートを受けており、ユタ州に拠点をもつモルモン家族研究と連結をもつ。完全家族と呼ばれるものにもとづく真の”RFLP factory”をサポートするのに年間200万ドル使われている。完全家族とは、4祖先、両親、少なくとも8人の子供からなる。HHMIは、また、1987年9月に開かれる、ヒトの遺伝子地図会議(No.9)の責任を引き受けるであろう。

Cahill氏は次ぎのように助言した。データベースに対して4-5年資金(fund)を出すことに加えて、NLMへゲートウェイされることが極自然で、普通の医者が必要な時にデータベースにアクセスできることである。HHMIでは、まだシーケンシングには手をだしていない。

日本 (この項、内田訳)

日本において幾つかの省または庁が生物学データベースに関心がある。大学研究者は文部省に依存している：その他に、厚生省、科学技術庁、農水省、通産省がある。

計算機ネットワークは日本では十分発達しておらず、また（生物学）データベースの所有権に関する考え方も（国際通念から）多少とも異なるようである。特に、DNAデータベースが国際的共有財産である一方で、民間財団－蛋白質研究奨励会－によって作られたデータベースはその財産であるとされている。

丸山によって説明されたように、文部省はDNAデータの入力作業を行う施設に予算を与えた。内田委員長であるDDBJ運営委員会には（この活動に関して）科学技術庁を代表する委員も、通産省から推薦された企業側委員も含まれている。DDBJの指し当たったの目標は日本で決定されたデータを入力することで、その量は全世界のデータの約10%になると推定される。DDBJは入力データの質および入力速度を許容できる程度まで改善することが必要であろう。内田の発言は他の発言者がふれた”1日百万塩基”までを含んでいない。

日本の各省庁は互いに競争関係にあるが、重複は互いに避けている。文部省がデータ入力を開始すれば、同じことを他の省庁はしない。JICSTは企業へのデータ配布活動を支援している。文部省は年額1/3US百万\$を5年間支出している。データの注釈は英文でされる。

現時点では和田教授が関係している”ヒューマン フロンティア計画”とは連結していない。

ヨーロッパ

L. Philipson氏はEMBO（欧州分子生物学研究機構）、EMBC（欧州分子生物学連盟conference）、EMBL（欧州分子生物研究所laboratory）の歴史と構造についてその概略を述べた。後者の最近の予算は年間2500万ドル相当で、260名の研究スタッフと総員510人を有する。EMBL核酸データベースに関しては、1982年の2名スタッフから現在の4-5名スタッフまで拡張した。コストは多分、NIHがGenBankデータベースに使ったおおよそ半分であろう。しかし、後者（NIH）は更にオンラインサービスなどを開始し、EMBLは注釈づけをより多くした。

この会議は、幾つかの異なったデータベースの必要性を示唆し、それらの間のアクセスの必要性、そしてソフトの開発の必要性などを示唆した、とPhilipson氏はうけとめた。

3年前、EMBLにおいて彼らはシーケンス、イメージ解析、グラフィクスを越えて、生物学計算(bio-computing)の理論の面の開発、タンパク質デザイン、シーケンスの相同性検索、ネットワークとデータ相互交換など幾つかの新しい面の開発の必要性

を感じとっていた。

現在の16人スタッフ（補助を加え、合計24人）の生物学計算スタッフから4年後には32人（合計50人）になることを申請している。Member Statesはこの拡張に要するコストの1/3をもつことを約束している。更に、1/3はBaden Wurttembergの州政府がもつであろう。残りの1/3をヨーロッパ経済共同体（EEC）に要求している。また、産業界からの援助も期待している。

EMBLはまた、Lee Hoodのそれとは違った手法の自動化したシーケンス技術の開発も行っている。EMBLは技術開発にも興味をもっている。

今回のような会議によって、協力、調整、データ管理などのセンターの位置づけが行われることは大変重要なことである。そして、EMBLは西ヨーロッパにおける、この役割をする最も自然な所であった。

K. Heusler氏は個人の資格ではあるが、このようなデータベースと基本的施設の根本的必要性を強調した。ヨーロッパの中で、種々な責任の所在をはっきりさす必要がある。化学工業界が作成した提案の中に、権限(authority)をもつヨーロッパの中心拠点(site)が必要であると強調している。同氏は更に、EMBLに一つの中心核をもつことは幸いなことだと感じた。彼が上げる必要性をリストすると、

- フォーマットとエントリーの内容をきめる。
- 電子的に入力されるデータの校閲システムを組織する。
- データベースをアップデートし、維持する。
- アメリカ、日本との国際連絡を保つ。
- 異なったデータベース間の相互参照をする。
- 他所に存在するデータベースへのアクセスサービス。
- 適切なメディアによるデータベースの配布。

適切な組織は適切な funding の必要条件である。正当な方向づけと基本的 funding が与えられてのみ、産業界を含む他の業界に寄与を要請することができる。利用者からの寄与は現金でもよし、或はデータ投稿の保証でもよい。

彼は更に、“ヨーロッパ生命情報サービス(European Bio-Information Service)”のアイデアを提案した。これはEMBLと密に連携しているが、別なもので、その運営委員にEMBL、EEC、政府、産業界を含める。その目的は、ヨーロッパの科学に生物情報(bio-information)を提供することである。

J. C. Thierry氏は2年前に作られた分子生物学における結晶学のヨーロッパ協会(European Association)について話した。その目的の幾つかを上げると、

- 一 コンピュータによるデータ、アクセス、トレーニングの確立に助言する。
- 一 結晶学における、検索、アレイ・プロセッサ、コンピュータグラフィック、データベース・マネージメントのためのソフトウェア開発の調整

EECの資金は2次元空間カウンターのための新ソフトウェア開発をサポートしている。彼らはまた、英国の"CCP4" (collaborative computing project) の教習を他の国へも広げた。

C. Saccone氏はイタリアの状況を説明した。イタリアにおけるデータベースの進展は、遺伝子工学とバイオテクノロジーの関連において、CNRからのモDESTな資金にたよっている。バリ大学の彼女自身の学科では、VAXコンピュータがあり、それはACNUC (核酸とタンパクシークエンスを選択や抜粋するソフト) やEMBLとGenBankを結合さすソフトなどを登載している。分子進化の計算のためには特別なソフトが開発されている。

彼女は情報の交換と調整が問題の主たるものと認識し、Grottaferrataワークショップを参照した。彼女は、どのようなことを国のレベルでサポートし、どのようなものを国際的サポートとすべきかを明白にし、同意を得ることが必要であると述べた。国のレベル、EECのレベル、EMBLのレベルと3重の支出に見えるようなことは批判されるであろう。役割をはっきり区別することが必要。

協力体制(collaborative mechanisms)

G. Hamm氏は、他の幾人かの講演者も指摘した、データのオーバーロードの危険に話題を戻した。このような状況に鑑み、データベースにおける優先順位を確立し、最低限界を設定する必要がある。それによって、止むを得なぬ場合、完全な崩壊というようなことを避けながら、その最低線まで下げることができる。K. Heusler氏が言ったように、データベースのセンターはその機能の最小限リストのようなものをもつべきである。そして、それはauthorityが決めるべきものである。Philipson氏が言ったように、安定した長期 funding によって次第に確立される信用によってのみ、authorityが達成され維持されるであろう。

NSFのJ. Dekens氏はデータベース・テクノロジーの標準(standards)に関し、鍵となる点に触れた。情報伝達のプロトコールにもとづく目的指向型システムが与えられたとして、何を標準化すべきかと言えば、それはアップデイト、アブストラクティングやその他の処理である。

DNAデータ入力：DDBJ Release 1.0

遺伝情報分析研究室
宮沢三造

ニュースレターNo. 6で報告したように、データ入力のための予算（データ入力委託費）が1986年度より付き、1986年12月から実際のデータ入力作業を進めてきた。サポートすべきDDBJ計算機も1987年3月から利用可能となり、データのリリースに向けてデータ入力システムの構築を進めてきた。その結果何とかエラーチェック用プログラムも稼働し、リリース可能なものが用意できるようになったので、1987年7月1.0版をリリースしGenBank, EMBLに配布した。66エントリー、約10万塩基である。（リリースノート参照）これは、現在のGenBank, EMBLと比較すると約1/100とまことに少ない。しかし彼等といえども、その1回目のリリースは同程度である。数は少ないが、まあまあの質のものが出来たと思う。データベースの作成は息の長い作業である。長い目で見ていただきたい。以下データ入力システムに関して、現状を述べてみたい。

DNA Data Bank of Japan

Release 1.0, July 1987

66 loci, 108970 bases, 4291 lines

This data base may be copied and redistributed freely,
without advance permission, provided that this
statement is reproduced with each copy.

Files included:

- 1) relnote1.0: this note
- 2) ddbj1.0: DNA data
- 3) ddbjd1.0: DNA protein coding sequences extracted
from ddbj1.0
- 4) ddbjpl.0: protein sequences translated from ddbjd1.0

Prepared by:

Takeo Maruyama General Manager
Sanzo Miyazawa Manager/Database Administrator
Hidenori Hayashida Scientific Reviewer
Motono Horie Secretary

The 3rd and 4th files above were generated by using
"seqext" and "pepttr" programs made by Dr. Jim Fickett
in GenBank, Los Alamos National Laboratory.

Acknowledgements:

We thank GenBank for helps, especially for providing us
with such tools that are useful for quality control in
data entry.

DNA Data Bank of Japan
National Institute of Genetics
Center for Genetic Information research
Laboratory of Genetic Information Analyses

1111 Yata
Mishima, Shizuoka 411
Japan

Phone: +81 559 75 0771 x647
E-mail: ddbj%niguts.nig.junet@relay.cs.net

LOCUS PHY300PLK 4870 bp ds-DNA entered 06/02/87
 DEFINITION Shuttle vector for Escherichia coli and Bacillus subtilis, plasmid
 PHY300PLK, complete sequence.
 ACCESSION N00054
 KEYWORDS PHY300PLK, shuttle vector, plasmid.
 SOURCE Plasmid PHY300PLK DNA, collected from the E.coli K12 strain C600.
 ORGANISM Plasmid PHY300PLK
 Artificial sequences; Cloning vehicles.
 REFERENCE 1 (base 1 to 4870)
 AUTHORS Ishiwa,H. and Shibahara-Sone,H.
 TITLE New shuttle vectors for Escherichia coli and Bacillus subtilis.
 IV. The nucleotide sequence of PHY300PLK and some properties in
 relation to transformation
 JOURNAL Jpn J Genet 61, 515-528 (1986)
 COMMENT PHY300PLK is one of the smallest hybrids of a new series of
 chimeric plasmids using the parental plasmids, pACYC177 of E.coli
 and pAM-alpha of Streptococcus faecalis. This shuttle vector for
 E.coli and B.subtilis contains an RNA primer gene for ori-177; the
 R-Tc gene; the R-Ap gene; Rep-alpha-1 gene, which is the plasmid
 replication gene; two replication origins, one for E.coli and the
 other for B. subtilis; and a polylinker. Twenty-three unique
 recognition restriction endonuclease sites have been found in
 PHY300PLK.

FEATURES	from	to/span	description
pept	1476	100 (c)	open reading frame (ORF) of tetracycline resistance (R-Tc) gene
pept	2687	1683 (c)	ORF of Rep-alpha-1 gene
pept	3289	4149	ORF of ampicillin resistance (R-Ap) gene
RNA	4350	4860	RNA primer for ori-177

SITES
 refnumbr 1 1 numbered 1 in [1]; zero not used
 site 1 42 a polylinker containing the recognition
 sequences for HindIII, XbaI, Bgl , PstI, AccI,
 HincII, Sall, BamHI, AvaI, SmaI, EcoRI
 ->pept 100 1 (c) ORF of R-Tc gene cds end (comp strand)
 pept<- 1476 1 (c) ORF of R-Tc gene cds start (comp strand)
 ->pept 1683 1 (c) ORF of Rep-alpha-1 gene cds end (comp strand)
 pept<- 2687 1 (c) ORF of Rep-alpha-1 gene cds start (comp
 strand)
 ->pept 3289 1 ORF of R-Ap gene cds start
 pept<- 4149 1 ORF of R-Ap gene cds end
 ->RNA 4350 1 RNA primer for ori-177 start
 RNA<- 4860 1 RNA primer for ori-177 end

BASE COUNT 1474 a 1026 c 867 g 1503 t
 ORIGIN the first base of the HindIII recognition sequence in the
 1 aagcttctag agatctgcag gtcgacggat ccccggaat tctgttata aaaaaaggat
 61 caattttgaa ctctctccca aagttgatcc cttaacgatt tagaaatccc tttgagaatg
 .
 .
 .
 4861 cgctttgccc
 //

Fig. DDBJ entry

データベース化の手順は 1) 作成システム (入力、校正、更新) 2) 検索システム 3) 解析システム に分けられよう。利用者の立場からは、特に 2) 3) が必要とされる。一方、一次データ作成を目的とするデータバンクにおいては、組織だったデータ収集、入力、更新を可能にする作成システムが欠かせない。これらのシステムはDBMSを用い一元的に管理するのが望ましい。しかし、作成に時間がかかりデータ入力をはじめているDDBJにはその余裕がない。我々は、作成が比較的容易であるため 3システムを別々に構築することにした。DBMSを用いる一元的な管理システムの構築は、並行して開発していくつもりである。

1. データ入力に関する国際協力：入力対象雑誌の分担

データ量の増加に伴い一機関ですべてのデータを網羅して収集することが困難になった。そのような背景から、GenBank, EMBL間での国際協力が叫ばれ、2,3年前より徐々にその協力が拡大されてきた。現在、データ入力に関しては、GenBank, EMBL間で学術雑誌の分担という形で協力がなされている。このような現状から、丸山 (遺伝情報研究センター長) が参加した1987年2月25日-28日ハイデルベルグの EMBL で "Future Database for Molecular Biology" と題する workshopにおいて、DDBJに対しGenBank, EMBLとの国際協力が要請された。データ入力に関しては、現在各データバンクのデータエントリーが雑誌からの入力である関係上、学術雑誌の分担が最も容易であるとの結論に達した。我々は、このような現状を考慮し、DDBJとしては日本で出版される学術雑誌を担当していくことにした。以下は、現在入力の対象にしている雑誌のリストである。(DDBJ Release 1.0 には、他の雑誌からの入力も含まれている。これは、雑誌の分担が決定する前に入力されたものである。)

Journals scanned by DDBJ

update: 12/25/87

Journals published in Japan:

Agricul Biol Chem	Vol. 50(01) ('86) - 51(09) ('87)	
Cell Struc Funct	Vol. 11(01) ('86) - 12(04) ('87)	
Chem Pharm Bull	Vol. 34(12) ('86) - 35(06) ('87)	
Devel Growth Diff	Vol. 28(01) ('86) - 29(05) ('87)	
J Biochem Tokyo	Vol. 100(01) ('86) - 102(06) ('87)	
Jpn J Cancer Res: GANN	Vol. 77(01) ('86) - 78(09) ('87)	
Jpn J Genet	Vol. 61(01) ('86) - 62(04) ('87)	
Microbiol Immunol		
Plant Cell Physiol	Vol. 28(01) ('87) - 28(05) ('87)	
Zool Sci	Vol. 3(01) ('86) - 4(05) ('87)	exc.4(03)

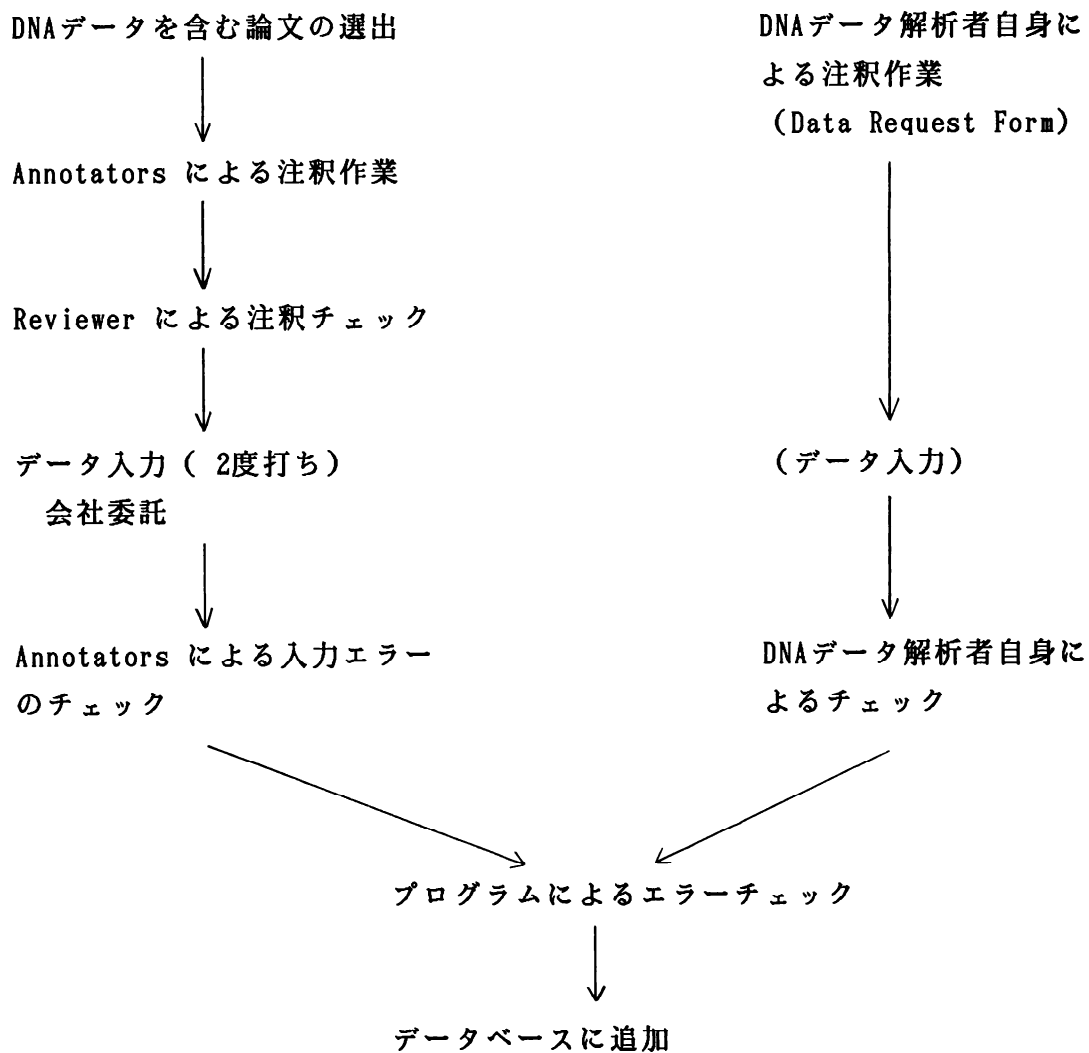
Journals published outside of Japan:

Immunogenetics	Vol. 26(03) ('87) - 27(01) ('88)
J Gen Microbiol	Vol. 133(01) ('87) - 133(11) ('87)
J Gen Virol	Vol. 68(03) ('87) - 68(11) ('87)
J Immunol	Vol. 138(01) ('87) - 139(10) ('87)

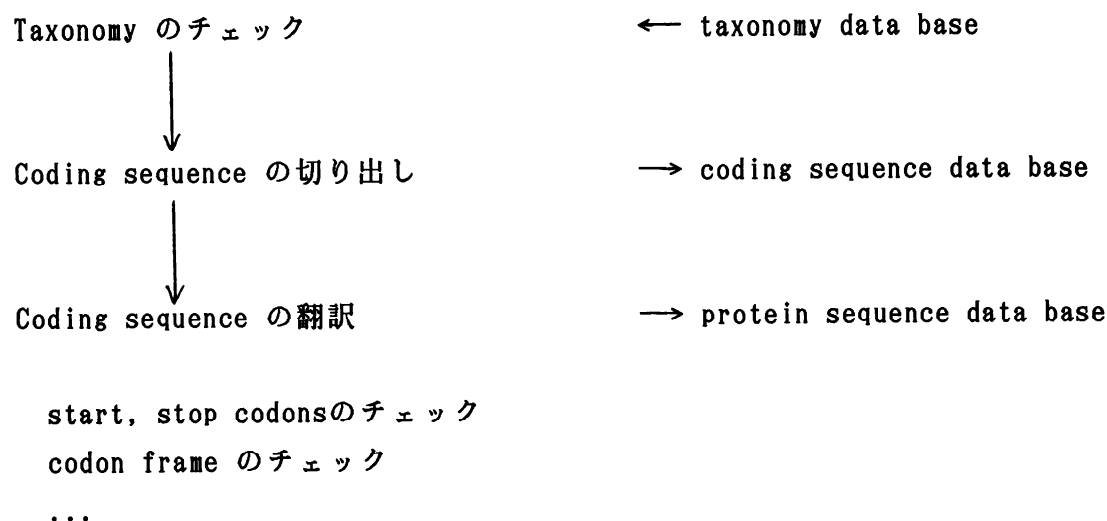
予想されるように、これらの雑誌に掲載されるDNAデータは数少ない。データの約80%を占めるJ. Biochem. まで、1年で数10編である。数少ないことは発足間もないDDBJにとっては有難いことではあるが、些か余裕があるため、外国雑誌を2, 3 入力の対象にしている。

データ入力の手順は以下のようである。

入力、校正プロセス



プログラムによるエラーチェック



ここでデータ作成システムの問題点とその改善について述べたい。データ作成システムの改善は一口で言えば、注釈作業の合理化とその支援システムの開発である。現在、注釈作業と注釈チェックはコーディング用紙にておこなっている。これは、主に 1) Annotator が計算機に不慣れなこと 2) データ入力用ソフトウェアが未開発のため、直接計算機入力が困難であることによる。プログラムによるエラーチェックも注釈者の手でなされるべきであり、そのためには注釈者への教育が不可欠である。現在注釈者の確保はそれ自体容易でなく、注釈者の確保と教育は、データ入力用ソフトウェアの開発とともに今後の課題である。

しかし、このような改善は、根本的解決策ではない。現在データバンクが直面している問題は、このような改善だけでは増大するデータに入力が追いつかないということである。又、完全な注釈をAnnotator に望むことは、不可能である。根本的解決策は実験家による注釈づけである。現在 DNA Data Submission Form (付属資料参照) を実験家に配布し、注釈づけをお願いしている。配布を効率的にするため、丸山が学術雑誌と話し合い、論文の掲載が決定した時点で、Data Submission Form を著者に送ってもらうよう、協力をお願いしている。下記の雑誌からは好意的な返事をもたらしているので、近いうちに実現されるであろう。皆さんの協力をお願いしたい。Data Submission Form の提出は、文書としてでも結構だが、計算機可読な形が望ましい。DDBJに連絡いただければData Submission Form が入っているフロッピーディスクを送付しますので、可能な限りフロッピーもしくは電子郵便で提出していただきたい。

協力要請した論文雑誌名

The Journal of Biochemistry

published by The Japanese Biochemical Society (monthly)

Microbiology and Immunology

published by Japanese Society for Bacteriology,
Society of Japanese Virologists, and
The Japanese Society for Immunology (monthly)

Japanese Journal of Cancer Research: GANN

published by Japanese Cancer Association (monthly)

Japanese Journal of Genetics

published by The Genetics Society of Japan (6 times a year)

Cell Structure and Function

published by

Zoological Science

published by The Zoological Society of Japan (bimonthly)

Development, Growth and Differentiation

published by The Japanese Society of Developmental Biologists
(6 times a year)

Plant and Cell Physiology

published by The Japanese Society of Plant Physiologists
(8 times a year)

Agricultural and Biological Chemistry

published by Agricultural chemical Society of Japan (monthly)

実験家自身の手によるデータ入力、今後の課題である。不慣れな人でも正しく入力できるような、すぐれたデータ入力用プログラムが必要とされよう。現在、このような問題も含め、データベースの望ましい改良に関し、EMBL, GenBank, DDBJ とで会合が開かれる予定である。

2. データ履歴管理

データ更新において最も大切なのは、データの履歴管理である。我々は現在UNIXオペレーティングシステムに付属する履歴管理システム、SCCS (Source Code Control System) を用い履歴管理を行っている。SCCSは本来プログラムの履歴管理が目的のため、DNAデータエントリーにはふさわしくない面もあるが、その利用の容易さから使用することにした。このシステムを使用することにより、

1) バージョン管理

任意なバージョンの作成、更新、回復

更新の理由、誰が更新したか、その日時

2) 排他管理

更新に関する排他性。

が可能である。このいずれも、複数の人によるデータ入力、更新の管理には、必須の条項である。

宮澤 三造

Nucleic Acid Research(NAR, vol.15, No.18, 1987)に掲載された "Announcement" 及び "A new system for direct submission of data to the nucleotide sequence data banks" の記事で既に御承知かと思いますが、NARとEMBLとの間で DNAデータの提出に関して協定が結ばれ、1988年 1月より実施されることになりました。その概略は

- Stage I : NAR への投稿者は、投稿前に DNAデータをEMBLへ提出し"accession number"を得ることが要求される。
- Stage II : EMBL (データバンク) 雑誌の出版時まで、提出された DNAデータを NAR の読者にオンラインでアクセス可能にすること。
- Stage III : EMBL (データバンク) は、提出された DNAデータを論文のreviewerにオンラインでアクセス可能にすること。

1988年 1月よりStage Iが実施され、その後Stage II、IIIへと移行が計画されています。

著者は以下のようなデータをデータバンクに提出することが要求されています。

a) 配列データ：計算機可読な形であること

著者が計算機を利用できない状況にある時のみハードコピーでの提出が許可される。

b) data submission formによる配列データの注釈づけ

c) 利用可能ならば、電子郵便の宛先、テレックス番号、テレファックス番号の通知

これらのデータは、当初EMBLへの提出が要求されます。EMBLは、データの受理後 7日以内にaccession numberを送付するか、提出データに不備な点があれば、再提出の通知をすることになっています。もちろん、7日以内の返答という条件は電子郵便、テレックス、テレファックスの時に通用されるもので、通常の郵便の場合は、その限りではありません。

データの提出先は、このシステムが順調に運営されるようになった時点で、他のデータバンク(GenBank等)にも拡大されることになっています。データバンク(EMBL, GenBank, DDBJ)としては、このNAR-EMBL間の協定を試金石と考え、他の雑誌とも同様な協定を結ぶよう努力する考えです。DDBJとしては、将来データ提出先として認められるよう準備していくつもりです。その間、著者はEMBLにデータを提出しなければならないわけですが、その際問題になるであろうことについて以下述べます。

まず、計算機可読な形でのデータの提出ということですが、近年パーソナルコンピュータの利用が盛んですので、配列データを入力している方も多いためその点問題はないかと思いますが、問題は媒体です。EMBLが取り扱えるものは

- (1) 電子郵便によるもの
- (2) 磁気テープ (9トラック)
- (3) フロッピーディスク：Macintosh もしくはIBM-PC用フロッピー (5.25")

です。フロッピーディスクの利用がまず頭に浮かびますが、御承知のように、フロッピー

ディスクは計算機のハードウェア、ソフトウェアに依存します。日本ではMacintosh, IBM-PC compatible の利用者は少なく、ほとんどはNEC-PC9801を使用のことと思います。IBM-PCで読めるよう、次ページの表を参照の上、フロッピーを format して下さい。MTの場合は問題ありませんが、MTの使用に慣れている方は少ないと思います。もう一つの手は電子郵便ですが、日本では残念なことに普及していません。DDBJの共同利用計算機は、現在電子郵便ネットワークは接続されています。データバンク間の国際協力の一貫としてDDBJは、EMBL, GenBankへのデータ提出等の目的のための電子郵便の利用は、利用登録をしなくても許可しております。どなたでも利用できます。内線電話、外線電話を利用しDDBJの計算機にパーソナルコンピューターを接続し、ddbnewsでloginし、データを転送しそのファイルを電子郵便でEMBL宛(emblsub@niguts.nig.junet)発送して下さい。DDBJは、そのメールをEMBL(datasubs@embl.bitnet)へ転送いたします。Data submission form も、オンライン用のものがDDBJ計算機から手に入ります。詳しくは、利用の手引を参考して下さい。

パーソナルコンピューターの接続にはモデムが必要です。費用がかかりますし、又ファイル転送、メール発送等慣れが必要ですが、数時間で届きますので最善の方法です。方法等、利用の手引に詳しく載っていますが、不明な点はお問い合わせ下さい。

なおファイルを作成する際には、どんなプログラムでも読めるよう以下のことに御注意下さい。

- 1) ファイルは単純なテキストファイルであること。
- 2) 一行の長さは 80 字以下であること。行末は、必ず改行で終わること。

特に、パーソナルコンピューターでデータを作成する際ワードスター等を用いるときは、必ずnondocumentとして作成して下さい。documentとしてファイルを作成するとワードプロセッシング用の文字が挿入され、他のシステムに転送した時意味をなさなくなりますので御注意下さい。

(注) PC9801で電子郵便を直接送受信するための最もよい方法は、ハードディスクが必要ですが、UNIX(PC-UX)を購入することです。UNIX計算機をお持ちの方は、0559-75-0771 内647 宮沢まで連絡下さい。

IBM-PC 5.25" or 3.5" compatible floppy disk format

Disk Characteristics	NEC PC-9801/XA/XL MS-DOS	IBM PC-XT/AT PC-DOS
2HD 1.2 MB = 1024 bytes/sector x 8 sectors/track x 77 tracks x 2 sides	Under version 2.xx or 3.xx >Format	cannot read/write
2HD 1.2 MB = 512 bytes/sector x 15 sectors/track x 80 tracks x 2 sides	Under version 3.xx or some ver. 2.xx >Format /5	can read/write, if 2HD can used.
2DD 720 kB = 512 bytes/sector x 9 sectors/track x 80 tracks x 2 sides	Under version 2.xx or 3.xx >Format /9	can read/write with ver. 3.30. In the case of other versions, it may not.
2DD 640 kB = 512 bytes/sector x 8 sectors/track x 80 tracks x 2 sides	Under version 2.xx or 3.xx >Format	can read/write with ver. 3.30. In the case of other versions, it may not.
2DD 360 kB or 2D = 512 bytes/sector x 9 sectors/track x 40 tracks x 2 sides	can only read	can read/write
2DD 320 kB or 2D = 512 bytes/sector x 8 sectors/track x 40 tracks x 2 sides	cannot read/write	can read/write with ver. 3.30. No data for other versions.

Footnotes:

- 1) 2HD means double side, high density and double track (96 TPI) floppy.
It is called double side and high capacity (2HC) in the U.S.A.
- 2) 2DD means double side, double density and double track (96 TPI) floppy.
- 3) 2D means double side, double density (48 TPI) floppy.

現在データバンク間で協議されている事項

宮沢三造

1987年11月データバンク間の会合が予定されている。予定されている主な議題は、

1) GenBank, EMBLフォーマットにおけるFeature/Site Tableの改良

利用者はよくわかると思うが、現在のFeature/Site Tableは、多くの欠点があり、その利用を限られたものにしてている。現在その大巾な改良に向けて作業が進行している。

2) Data submission form の作成

DDBJ/EMBL/GenBankで共通のData submission formを使用するという前提で、その統一案を作成する予定である。

3) データ記述言語の標準化とデータ入力フォーマットの記述

現在、データバンクでは、データベースをFlat file の形で管理している。データベースの更新、追加、データベースの検索機能がよりすぐれたデータベース、データ作成、検索システムを一元的に管理するようなデータベースを構築するのが望まれる。例えば、Relational data base(RDB) の使用である。

しかし、RDB を用いた時、現在のようなファイルフォーマットによる配布だけでなく、データベースを完全な形で配布可能な、又、データベースの変更をも容易であるような配布形態が望ましい。データベースの記述をデータベースの一部として含むようなデータ配布である。データベースの記述は、計算機で可読であるだけでなく処理可能なものでなくてはならない。そのためには記述（言語）に関して標準化が必要である。

このようなデータ記述法に関する標準化は、又実験家によるデータ入力の際にも効力を発揮する。このような標準化された記述法でデータ入力フォーマットを書くことにより、入力フォーマットの変更に影響されないような入力用ソフトウェアが比較的容易に作成可能となる。

会合では、標準化案、ならびにデータ入力フォーマット、データベース構造について議論される予定である。

データバンク間での協議連絡は、現在電子郵便でおこなっている。参考のために、データバンク間での連絡用 Mailing list を次ページに示す。

Mailing lists for electronic mail communication among data banks
(Taken from a mail sent from benton@bionet-20.arpa in 10/16/87)

Dear data-bankers:

The following mailing lists have been established on BIONET-20.ARPA: nucall, ddbj, embl, gb, gb-ig, gb-lanl, gb-nih, and genbank. The genbank address is intended for queries from the user community. The other addresses are for reaching the staffs of the three cooperating data banks and are related as diagrammed below. (note that the gb-lanl address will soon be changed to point to a mailing list maintained on a machine at LANL). If no domain is given, the address is @BIONET-20.ARPA.

Thank you for responding to the testing this morning. I think all the addresses work.

-db

```
nucall -|- ddbj -- ddbj%nigsun.nig.junet@relay.cs.net
          |          | smiyazaw@niguts.nig.junet      (Sanzo Miyazawa)
          |          | hhayashi@niguts.nig.junet      (Hidenori Hayashida)
          |          | tmaruyam@niguts.nig.junet      (Takeo Maruyama)
          |
          | - embl -- DataLib@EMBL.bitnet
          |          | CAMERON@EMBL.BITNET      (currently Graham Cameron,
          |          | KAHN@EMBL.BITNET          (Patricia Kahn,
          |          | HAZLEDINE@EMBL.BITNET      (Dave Hazledine,
          |          |                               (Shirley Jones, and
          |          |                               (Guenter Stoesser)
          |
          | - gb --
          |   - gb-ig -----| benton@bionet-20.arpa  (David Benton)
          |                   | kelly@bionet-20.arpa  (Mike Kelly)
          |                   | rob@presto.ig.com      (Rob Liebschutz)
          |                   | brutlag@bionet-20.arpa (Doug Brutlag)
          |                   | bossinger@bionet-20.arpa (June Bossinger)
          |
          |   - gb-lanl ----| cb@lanl.gov          (Christian Burks)
          |                   | ct@lanl.gov          (Chang-Shung Tung)
          |                   | dan@lanl.gov          (Debi Nelson)
          |                   | dd@lanl.gov          (Dan Davison)
          |                   | fxxm@lanl.gov        (Frances Martinez)
          |                   | jwf@lanl.gov        (Jim Fickett)
          |                   | kam@lanl.gov        (Kersti MacInnes)
          |                   | kms@lanl.gov        (Karl Sirotkin)
          |                   | mxxs@lanl.gov        (Maura Smith)
          |                   | skm@lanl.gov        (Santosh Mishra)
          |                   | tgm@lanl.gov        (Tom Marr)
          |                   | wbg@lanl.gov        (Walter Goad)
          |                   | fkh@life.lanl.gov    (Kay Houghton)
          |
          |   - gb-nih ----| czj@nihcu.bitnet    (Jim Cassatt)
          |                   | jp2@nihcu.bitnet    (Jane Peterson)
          |                   | dab@vax2.nlm.nih.gov  (Dennis Benson)
          |
genbank -----| benton@bionet-20.arpa
          |          | kelly@bionet-20.arpa
          |          | brutlag@bionet-20.arpa
          |          | kristofferson@bionet-20.arpa (Dave Kristofferson)
          |          | cb@lanl.gov
          |          | tgm@lanl.gov
```

長らくお待たせいたしました。外線電話、DDX-パケット回線が敷設され、また DDBJ 計算機システムの所外利用に関する利用規定も定まりましたので、計算機のオンライン使用を1987年 9月より開始しています。営利機関、非営利機関の別なく利用可能です。

なお計算機を利用する場合は以下の機材、手続きが必要です。(利用の手引参照)

- 1) 端末、もしくはパーソナルコンピュータと端末エミュレータープログラム。
NEC-9800用端末エミュレータープログラム(VTエミュレーター)とデータ転送端末エミュレータープログラム(Kermit)を希望者には配布しますので、申し込み書をお送りください。
- 2) モデム
全二重モデムなら何でも結構です。通信速度の速い(2400 bauds)ものでできればエラー修正をする MNP モデムを推薦します。利用の手引を参考にしてください。
- 3) 端末とモデムを接続する RS232C ケーブル
パーソナルコンピュータの付属品として購入できます。
- 4) 外線、不可能な場合は内線電話(モデムを接続するため電話線と受話器との接続をモジュラー方式に変換すること。部品を購入すれば個人で簡単にできます。) 遠隔地ではパケット通信を利用するのが割安かも知れません。外線電話を利用する第二種パケット通信回線の申請例は利用の手引を御覧下さい。
- 5) DDBJ 計算機利用申請書提出
当面試験利用ですので、利用料金は無料です。

営利機関、非営利機関の別なく研究者であれば利用できますのでどうぞご利用下さい。

DDBJ 計算機システムは、UNIXの電子郵便、電子掲示板ネットワーク JUNETに加入いたしました。現在、データバンク間(DDBJ, GenBank, EMBL)の連絡は全て電子郵便でおこなっています。DDBJ, GenBank, EMBL へのデータサブミッションに電子郵便をご使用下さい。当然計算機可読ですので、GenBank, EMBLもそれを望んでおります。現在、外国向け電子郵便の発着信は有料となっていますが、GenBank, EMBLへのデータ送付及び連絡は、DDBJが自動的に転送します。電子郵便は数時間で到着しますので、非常に便利です。国内、外の研究者との研究連絡用としてもご利用下さい。詳しくは利用の手引を参照下さい。

DNA データ送付用ならびにDDBJ活動に関する広報のために、特別のアカウント(ddbjnews)を設けました。詳しい情報を得るには、計算機にログイン後 "getinfo"コマンドを利用して下さい。又、このコマンドを使用すると、システム及びデータベースに関する情報はもとより、データ配布、ソフトウェア申し込み用紙等も手に入ります。申し込み用紙にエディターを用い記入の上、ddbjあて電子郵便で申請書を送り下されば、それに従い処理致します。DDBJ側の労力を減らすためにも、このような事務にも計算機を利用下さるようお願い致します。御意見、質問等もメールの形でお寄せ下されば助かります。計算機へのアクセスの方法は、利用の手引を参照下さい。

計算機導入から日が浅く、UNIXシステムでは DNAデータ解析用プログラムは、今後、順次開発する予定で、現在では十分ではありません。不十分な面は、当面 VAX/VMSを使い、補って頂くようお願い致します。M380Q/UTS とVAX/VMS のほとんどのプログラムに関してオンラインマニュアルが利用できます。今後印刷物としての配布も考えていきたいと思っています。

現在 DDBJ に収集されている配布可能な核酸および蛋白質データベースは以下のようである。PGtran は GenBank 35 版からの翻訳データベースである。(Claverie et al., Nature 318, p19, 1985) SWISSPROTはEMBLフォーマットを用いる蛋白質データベースでそのほとんどは PIRからの変換である。しかし独自に入力も行っている。その他、レトロウイルスデータベース、免疫関連DNA、蛋白質データベース等利用可能なものがある。(利用可能なデータベースのリストは、“getinfo” コマンドを利用して得られる。)

DNA 塩基配列データ：

DDBJ	1 版 (07/87)	66 エントリー	108,970 塩基
EMBL	13 版 (10/87)	14,397 エントリー	16,023,442 塩基
GenBank	50.0 版 (05/87)	12,534 エントリー	13,048,473 塩基
NBRF	31.0 版 (06/87)	2,288 エントリー	4,711,652 塩基

蛋白質アミノ酸配列データ：

PIR	13.0 版 (06/87)	4,525 蛋白質	1,116,951 残基
PGtrans	35.0 版 (09/85)	3,107 蛋白質	653,339 残基
SWISSPROT	5 版 (09/87)	5,205 蛋白質	1,327,683 残基

配布媒体は GenBank の場合は磁気テープとフロッピーディスク、その他は磁気テープのみである。GenBank フロッピーディスクは MS-DOS 版であり、GenBank フロッピーディスクに付属して配布される IBM-PC 用の検索プログラムは、NEC-PC9800 用に移植したのでデータと共に配布可能である。全データは、GenBank フロッピーディスクが 2DD (720 kB format) フロッピーで 13 枚そして NEC-PC9800 用の検索プログラムその他が 3 枚の 16 枚である。磁気テープのフォーマット等については資料として付属の申し込み書を参照願いたい。すべて配布媒体はコピー後送り返すことをお願いしている。配布は申し込み書をお送りいただくから 2, 3 日以内、定期配布希望者への新版の配布は新版到着後 1 週間以内に発送している。

以下は各データベースの簡単な収集内容である。なお、図はこれまでの各データベースの収集数の変遷を描いたものである。EMBLの収集塩基数が1985年中頃急増しているが、これはGenBankと相互にデータ交換を始めたためと思われる。

1. GenBank Release 50.0

Group	#Entries	#Bases
Primate	1565	1930507
Rodent	1792	1713598
Other mammalian	372	394441
Other vertebrate	545	518518
Invertebrate	775	676378
Plant	782	1013120
Organelle	449	881664
Bacterial	1027	1482299
Structural RNA	687	76128
Viral	1211	1831351
Phage	179	296959
Synthetic	253	79232
Unannotated	2897	2154278

2. EMBL Release 13

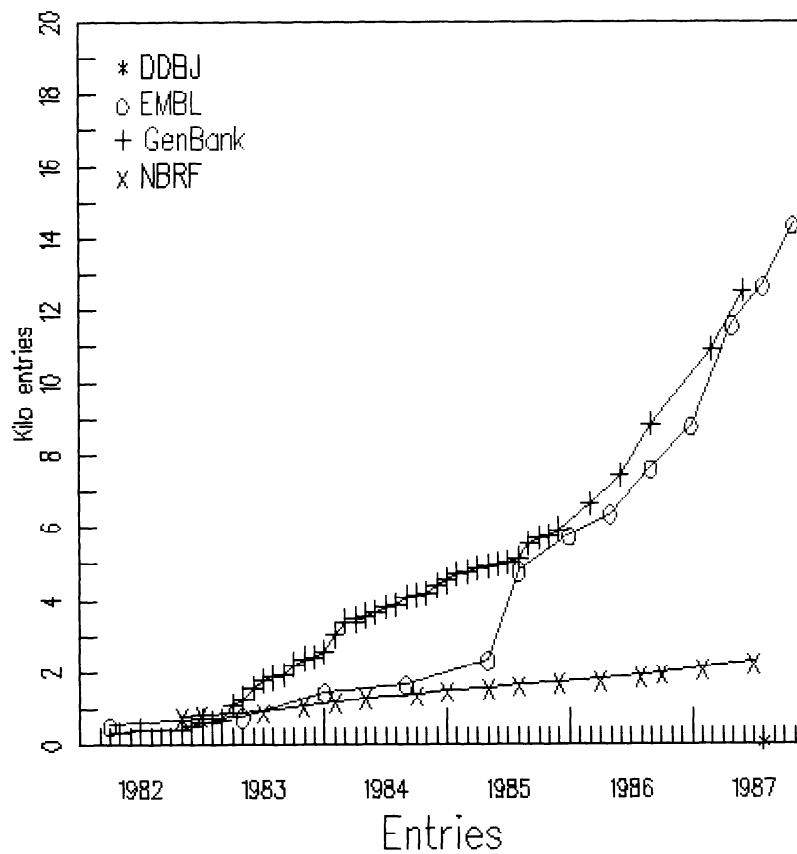
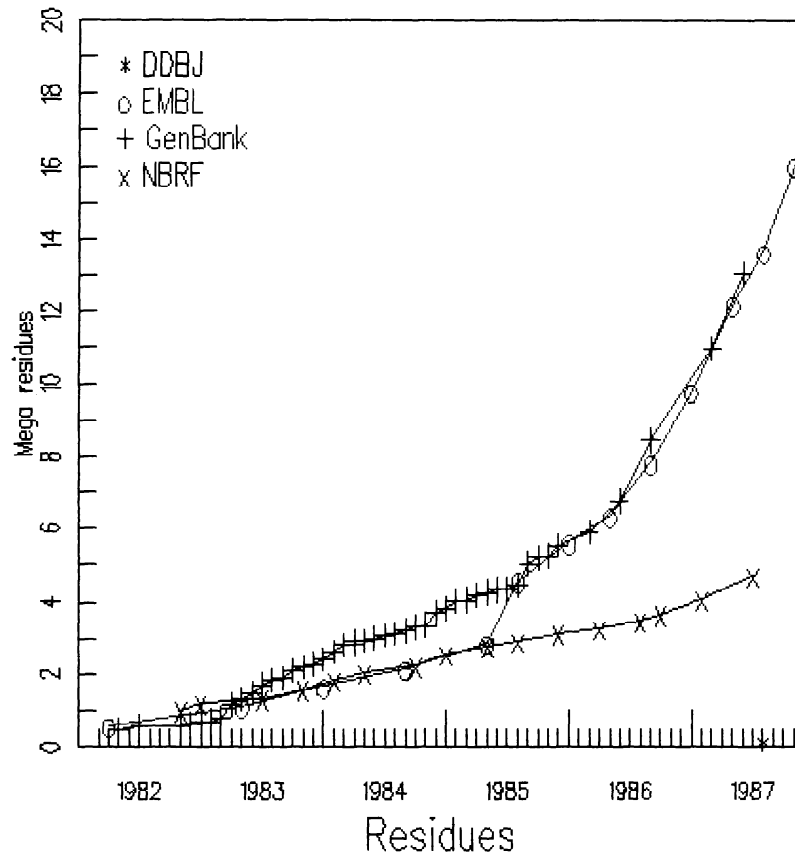
Group	#Entries	#Bases
Artificial	223	91717
Chloroplast	203	520683
Genetic elements	58	51617
Mitochondrial	367	434795
Prokaryotic	1501	1846153
Viral/Phage	1567	2657169
Eukaryotic	7467	8133595
Unclassified	25	48744
Unannotated	2986	2239005

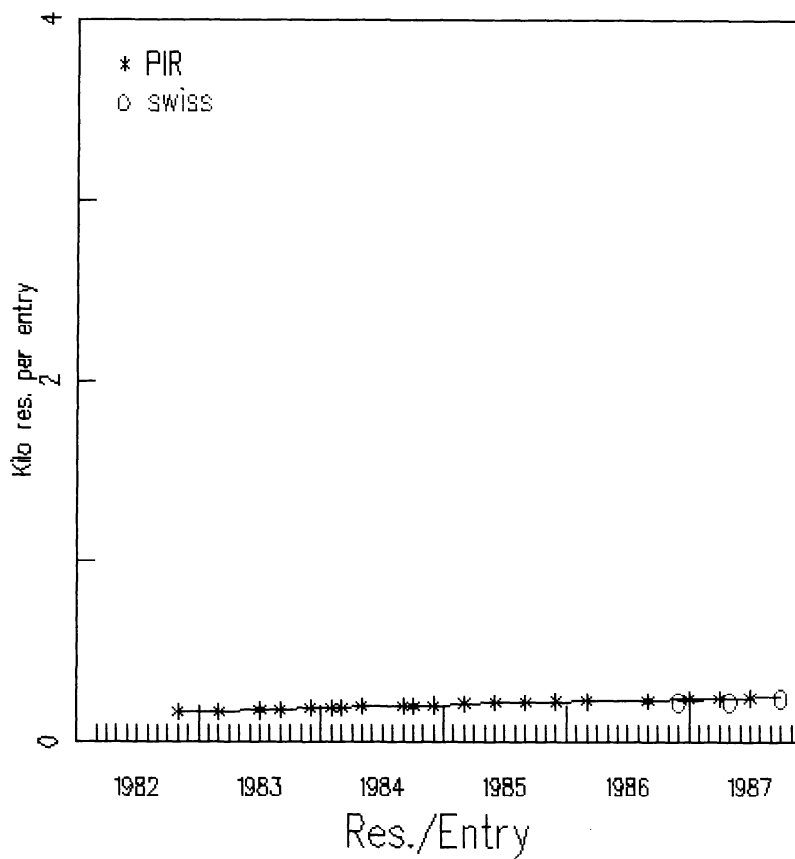
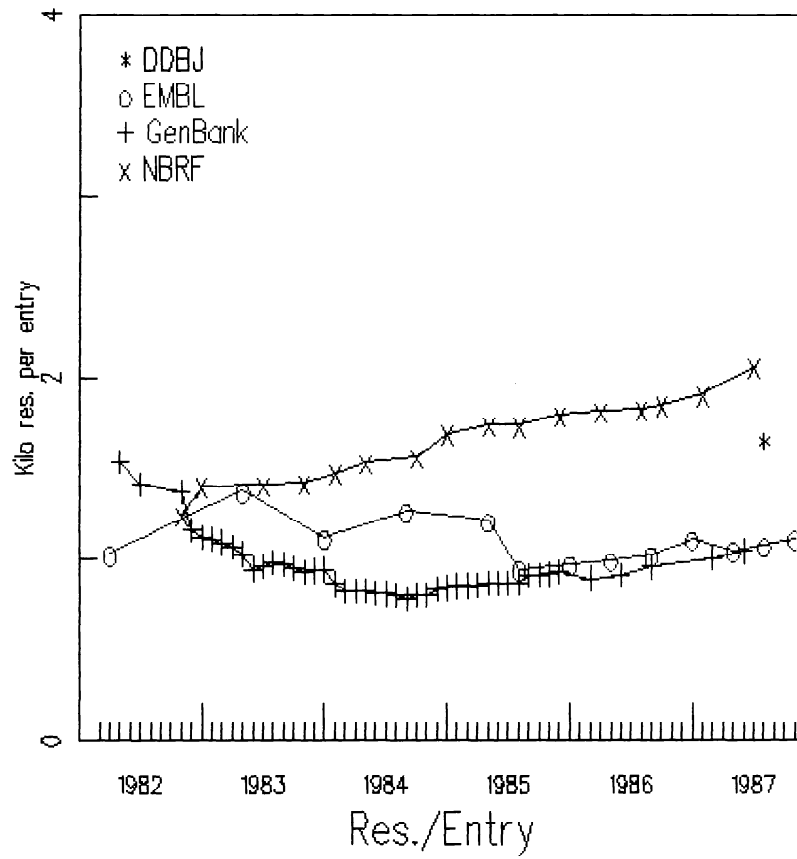
3. NBRF Release 31.0

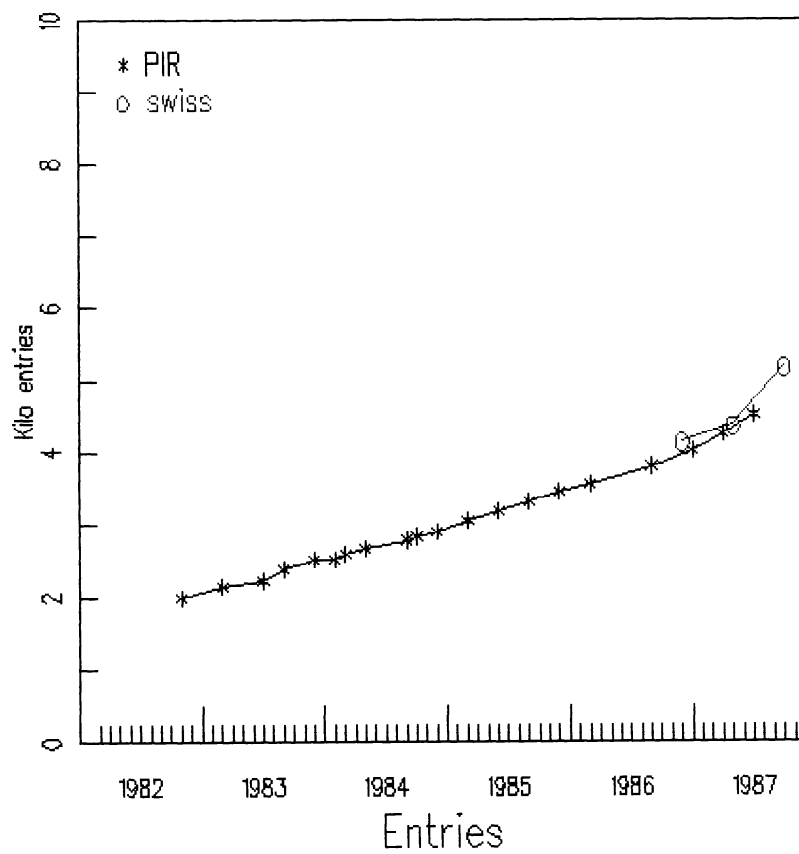
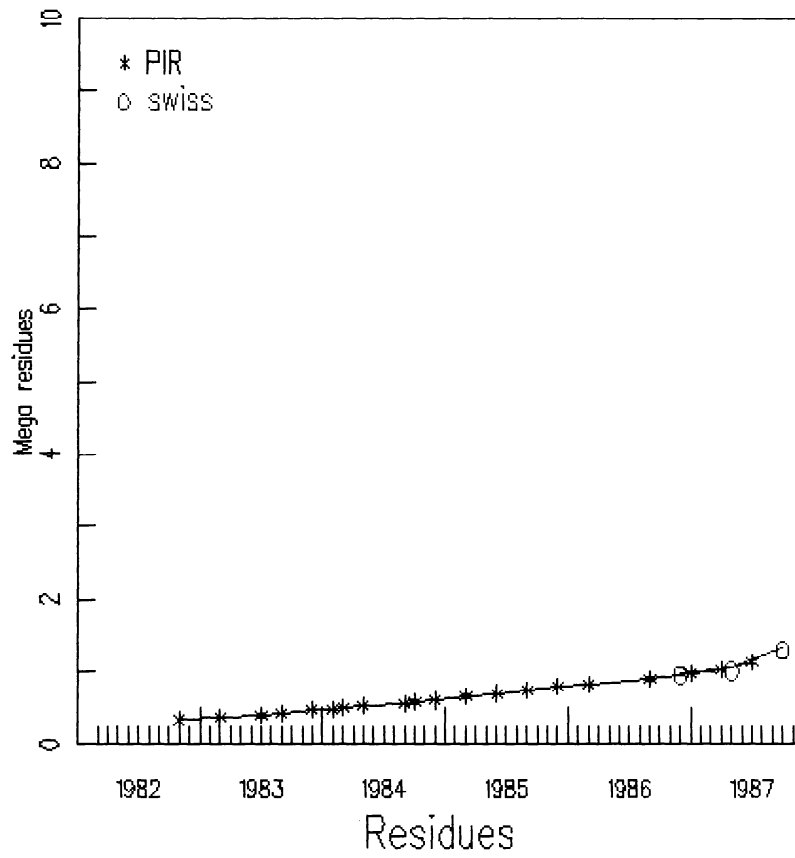
Group	#Entries	#Bases
Eukaryotes	2288	2449858
Mammals	779	1396208
Plants and fungi	259	637221
Eukaryotic viruses	400	1352139
Prokaryotes	460	694144
Bacteriophages	64	215511
Animal viruses	351	1221440
Plant viruses	49	130699
E. Coli	225	382423
Fungi	176	259209
Human	285	658168
Mitochondria	61	200115
Chloroplasts	30	31995

4. PIR Release 13.0

Group	#Entries	#Amino acids
Eukaryotes	2841	579525
Mammals	1575	342985
Plants	327	62610
Fungi	157	48262
Prokaryotes	741	167591
Animal viruses	625	284818
Plant viruses	67	28961
Bacteriophages	252	56754







DNA データベースの配布に関する活動報告

DDBJ 堀江 元乃

米国からGenBank, NBRFデータベース、欧州からEMBLデータベースを磁気テープで取り寄せ、希望者に配布している。その他蛋白質データベース（NBRF-PIRデータベース、GenBank 35版からの翻訳データベース PGtran）も希望者には配布している。配布媒体は GenBankの場合は磁気テープとフロッピーディスク、その他は磁気テープのみである。GenBankフロッピーディスクに付属して配布されるIBM-PC用の検索プログラムは、NEC-PC9800用に移植しデータと共に配布した。配布形態は定期もしくはは一次配布である。磁気テープの配布総数は 580本である。フロッピーディスクの配布枚数は 746枚（約16枚/件）である。今年度の配布実績の詳細は以下のようである。

DDBJニュースレター	大学/研究所	営利企業	合計
No.6 配布数	792	153	845
定期購読希望者数(87/11/07)	116	23	139
定期配布希望部数	134	29	163
計算機利用申請書郵送数	198	13	211
計算機所外利用者数	19	6	25
Kermitプログラム配布数	14	3	17

GenBank (87/10/31 現在)

版		定期配布		一時配布		合計	
		大学	企業	大学	企業	大学	企業
40	86/02	1	0	0	0	1	0
42	86/05	6	4	1	1	7	5
44	86/08	13	10	3	1	16	11
48	87/02	10	10	2	1	12	11
50	87/05	19	11	4	2	23	13

EMBL (87/11/05 現在)

版		定期配布		一時配布		合 計	
		大学	企業	大学	企業	大学	企業
8	86/04	6	5	3	2	9	7
9	86/09	10	8	1	0	11	8
10	86/12	10	8	0	0	10	8
11	87/04	12	8	3	5	15	13
12	87/07	11	8	0	0	11	8
13	87/10	15	10	0	0	15	10

NBRF (87/10/31 現在)

版		定期配布		一時配布		合 計	
		大学	企業	大学	企業	大学	企業
27	86/03	4	3	2	0	6	3
28	86/07	5	5	2	1	7	6
29	86/09	8	7	1	0	9	7
30	87/01	8	7	0	0	8	7
31	87/06	10	5	0	0	10	5

NBRF VAX/VMS版 (87/10/31 現在)

版		定期配布		一時配布		合 計	
		大学	企業	大学	企業	大学	企業
28	86/07	0	1	0	0	0	1
29	86/09	2	2	1	3	3	5
30	87/01	2	2	2	3	4	5
31	87/06	2	4	0	1	2	5

PIR (87/10/31 現在)

版		定期配布		一時配布		合計	
		大学	企業	大学	企業	大学	企業
8	86/02	9	6	2	2	11	8
10	86/08	12	8	0	0	12	8
11	86/12	13	8	0	1	13	9
12	87/03	16	6	2	1	18	7
13	87/06	16	6	0	1	16	7

PIR VAX/VMS版 (87/09/03 現在)

版		定期配布		一時配布		合計	
		大学	企業	大学	企業	大学	企業
8	86/02	0	1	0	0	0	1
9	86/05	0	1	0	0	0	1
10	86/08	3	4	0	1	3	5
11	86/12	4	5	0	0	4	5
12	87/03	4	7	0	1	4	8
13	87/06	3	8	0	0	3	8

PGtran (87/04/30 現在)

版		定期配布				一時配布				合計	
		大学		企業		大学		企業		大学	企業
		9	2	4	0	1	1	1	0	10	3
35	85/09	11		4		2		1		13	5

GenBank 圧縮版 フロッピー (87/11/05 現在)

版		定期配布		一時配布		合計	
		大学	企業	大学	企業	大学	企業
40	86/02	1	0	1	0	2	0
44	86/08	10	0	28	5	38	5
48	87/03	0	0	1	1	1	1

データバンク利用者へのお願い

1. Data Request Form の提出は、文書としてでも結構だが、計算機可読な形が望ましい。DDBJに連絡いただければData Request Form が入っているフロッピーディスクを送付しますので、可能な限りフロッピーもしくは電子郵便で提出していただきたい。
2. DDBJ計算機のオンライン利用が可能になりましたので、データバンクへの質問、書類請求、送付等、可能な場合は、オンラインで電子郵便を利用してください。
3. データの磁気テープ配布に関し、テープ装置のドライバー（ソフトウェア）にバグがあり、利用者に迷惑をおかけしました。今後もおかしな点がありましたら、御連絡ください。
4. データベース利用者へ ---- 英小文字使用について

GenBank, EMBL, NBRF, NBRF-PIR のオリジナルなデータベースは全て英小文字を使用している。また読み誤りを防ぐため、塩基も従来からの大文字表現 (A,T,C,G) から小文字表現 (a,t,c,g) へと移行しつつある。小文字表現は IUB (International Union of Biochemistry) でも認められ、GenBankは1985年 8月 35版は既に塩基を小文字で表現している。又 EMBLも小文字表現に移行しつつある。このような現状を考えると、利用者も小文字表現を許すようプログラムの変更することが強く望まれる。

なお、配布したデータ (EBCDICコードを用いたもの) がカタカナで表示されたのでデータが破壊されているのではないかとの苦情がよく寄せられる。これはIBM 互換機 (FACOM) で EBCDIC コードの小文字をカタカナに使用する 変形 EBCDIC コードをプリンター、ディスプレイで使用しているからにすぎず、データエラーではない。研究機関でカナの使用はまず考えられないからコード体系として小文字を含む EBCDIC をプリンター、ディスプレイ等で使用することをお勧めする。変換は容易で SE にプリンター、ディスプレイの文字フォントの入れ換え (通常フロッピーの入れ換え) を依頼すればよい。費用はかからないので是非変更することをお勧めする。なお EBCDIC コードは特殊文字に関して、IBM, FACOM, HITACで異なる。可能なかぎり ASCII コードで請求なさることをお勧めする。

(宮沢三造)

5. 配布磁気テープの返却のお願い --- 特に大学関係者へ

データ配布のための磁気テープが返却されないケースが大学関係者に多く、迷惑しています。配布した磁気テープはできるだけ早くコピーして返却して下さいようお願いいたします。

(堀江元乃)

DDBJ アンケート

ニューズレターNo.6でアンケートをお願い致しましたが、多くの方から返答をいただき有難うございました。今後のデータバンク運営に生かしたいと思います。
参考のため、アンケートのまとめを以下に載せます。

結果のまとめ

堀江 元乃

回答数 113 (大学・研究所 94 / 企業 19)

研究分野／業務は、DNA データベースに関係する分野ですか？

YES 93 (76 / 17) NO 20 (18 / 2)

研究分野／業務内容

非営利機関 (大学・研究所等)

分子細胞生物学

遺伝子の複製とその制御など

大腸菌を中心としたグルタチオン生合成系酵素遺伝子の発現・構造解析

好熱性細菌の糖質関連酵素遺伝子の解析

遺伝子の構造解析および蛋白質のアミノ酸配列の研究への応用

生体分子の構造や性質を明らかにする

発癌物質の代謝酵素の研究

腫瘍マーカーに関する酵素のクローニング

分子生物学

Gene cloning / DNA sequenceing

新しいがん遺伝子の構造解析など

動物細胞の分子遺伝学の研究

細菌毒素遺伝子の研究

情報処理の方法論についての研究

分子遺伝学・分子進化学

突然変異によって生じた遺伝子の塩基配列変化

腫瘍ウイルス学・分子生物学

高等植物用ベクターの開発、高等植物の形質転換

葉緑体遺伝子の解析

葉緑体DNA の構造と機能に関する研究

cDNA cloning

腫瘍遺伝子の構造解析

癌の分子生物学

リンフォカイン遺伝子に関する研究

蛋白質工学・生物分析化学・応用微生物

酵素プラスミドの構造と機能

視物質

高度好塩菌のレチナール蛋白質

外来遺伝子の動物細胞への移入

移入遺伝子の動物細胞よりの回収と構造決定

生体エネルギー変換タンパクの構造と機能、

生体エネルギー変換タンパクの比較生体学と進化

植物における遺伝子のクローニング

カビのクローニング等

グラム陽性細菌の遺伝子発現

真核細胞のリボソーム蛋白質の構造と機能

免疫系細胞特異的遺伝子の解析

高等生物遺伝子の構造、発現、機能 分子医学

真核生物の分子生物学

分子遺伝学

遺伝情報の発現の調整制御機構

細胞増殖を制御する遺伝子の研究

薬理学、神経薬理学

真核細胞のタンパク質分泌にあずかる遺伝子の研究

ヒト胚細胞腫瘍を用いた分化の研究

腫瘍細胞表面抗原発現の分子機構に関する研究

急性リンパ芽球性白血病関連抗原の遺伝子クローニング

マウス主要組織適合抗原遺伝子の発現の解析

肝炎ウイルス遺伝子の解析

タンパク質の部分構造と機能、細胞内局在など

蛋白質・核酸の物理化学

分子生物学研究

Cell biology / genetics

大腸菌および酵母のリボソーム生合成の分子機構

ヒト遺伝病の解析

ヒト遺伝病遺伝子の解析

レトロウイルス、ハシカウイルス、HFRSウイルスの形態学的研究

植物分子生物学

ウイルス学

分子生物学／DNA複製、遺伝子の構造、機能解析

細胞生物学、分子生物学、放射線生物学

サラセミア患者の遺伝子解析

細胞生物学、 μ VAXII System management

分子進化及び蛋白の構造・機能

核酸および核酸・蛋白質複合体の分子構造研究

ペプチド・蛋白質の機能に関する理論的研究

蛋白質・核酸の立体構造の解明、蛋白質・核酸のデータベースの整備と利用

分子生物学、プロテインエンジニアリング

蛋白質の高次構造と機能の予測

蛋白質の立体構造解析

遺伝子、蛋白質工学

DNA の高次構造

リン脂質合成酵素遺伝子の単離と構造解析

RNA 高次構造の解析、タンパク高次構造の予想

蛋白質の構造と機能に関する研究

ホモロジー検索へのデータベースの利用

sequence analysis

DNA ・蛋白質一次配列の解析プログラムの開発・応用

蛋白質・核酸配列データのコンピューター解析

遺伝子発現に関するシグナルの共通配列の検索

遺伝子の構造解析

5S RNAの一次配列とそれにもとづく系統樹の作成

葉緑体遺伝子構造と他のものとのホモロジー検索

ミトコンドリア蛋白構造と他のものとのホモロジー検索

蛋白質配列データベースの作成

蛋白研の大型計算機にデータベースを構築すること

DNA データベースを一般ユーザーに提供する

ペプチドの文献についての抄録誌発行

ファクトデータベースの開発、稼働

数理生物的な仕事、学内共同利用計算機の管理業務

営利機関（企業等）

様々な酵素の研究

遺伝子操作

データベース管理

微生物、動・植物解析

発現ベクター作製

発現ベクター改良

遺伝子工学を利用した分子生物学

細胞生物学

蛋白質工学

遺伝子工学

生化学

HLA の解析

分子遺伝学

遺伝子クローニング

蛋白質の分離・精製

遺伝子検索、蛋白質解析、遺伝子発現

微生物変換、それに関する遺伝子のクローニング及びその発現

遺伝子のクローニング及びその発現

動物細胞の培養、DNA アミノ酸配列の解析 データベースの管理

遺伝情報解析の研究

パソコン用核酸・アミノ酸配列解析ソフトの開発

データ解析、データベース管理、解析システムの開発・導入

DNA ・蛋白質解析ソフトウェアの開発研究

分子設計、ドラッグデザイン及び計算機システム管理

YESの場合、DNA配列決定のような実験に関係していますか？

YES 76 (66 / 10) NO 17 (10 / 7)

YESの場合

DDBJでは、論文雑誌との協力の下、DNA配列決定の論文は accept された時点で、著者にデータベース登録用のデータ記入をお願いすることを計画しています。協力していただけますか？

YES 74 (64 / 10) NO 2 (2 / 0)

NO の場合、その理由

- 印刷になるまでに自分の研究を進展させるため
- 雑誌社からの提出に任せたい。

DNA配列収集についての意見

- 日本でも独自にすべきである
- documentation等が必要最小限のものでよいから、速報版を出してほしい
- 現在のままでよい
- PC-9801用の編集エディターを配って欲しい
- 良いことです
- 重要かつ有益である
- 特にGenBankと友好的かつtightな関係を将来にわたって維持してほしい
- 検索の仕方が良いものがよい
- 得られた配列に入力ミスがないよう二重にチェックする
- 各学会の協力が必要

論文投稿者自身の段階での入力（記載）ミスをできるだけ少なくするようなことが重要

- やる以上はGenBank, EMBL, NBRFに並ぶ立派なものになってほしい
- GenBankやEMBLとの重複は無駄。CAS on-line等でも引用できるのでは
- あらゆる形のディスクを送付できるとありがたい
- 新しいDNA 配列を早く利用できるようにしてほしい
- アルファベットの印刷体を機械がデータとして読み取るのは容易なはずだから、メーカーに装置開発を急がせて手入力、音声入力はやらなくてもすむようにした方がよい
- 大変便利で有効価値が大。on-line でつなげたい
- GenBank あるいはEMBLのような大規模なdata base はあちこちで作る必要はなく、むしろそれを完全なものにする国際的な協議実施機関ができればよい
- AIDSウイルス関係を多く集めてほしい
- Up to dateなData Base になるようなシステムを組んでほしい
- sequence data の急増にどうやって対応していくかが問題
- なるべく早い段階に著者の協力を定着させることがよい
- 国内、国外を問わず、なるべく早い収集を希望
- DNA 塩基配列だけでなく、アミノ酸配列に関しても同様なデータ入力を同時に実施するようにした方がよい

DNAデータベースの format ならびに annotationについての要望

- GenBank format だけでなくEMBL formatも平等に扱ってほしい
- ニュースレターに記載の通りでよい
- 従来のformatと互換があるものがよい。新しいformatの場合、変換program をsupport してほしい
- EMBLのように全部を一つのfileに入れると検索に時間がかかるので、何項目かに分けてデータベースとしてほしい
- EMBL format を accept してほしい
- 全てのデータベースのformatが統一されることが望ましいが、少なくとも相互交換が容易にできるようにしてほしい
- sequenceの60ケタごとの頭にある数値をはずしてほしい
- formatは利用しやすい形態を望む
annotationはできるだけ詳しいことを望む(DNAデータベースは配列の方を重視し、annotationを軽くみているように思う)
類似、同一、別名の配列データの扱いをどうするかが一つの問題になるであろう
- GenBankのformatにするのはよいことだ
- PC-9801用の編集エディターに、あらかじめ標準的format table の記入事項を組み込み、それに答える形にする
- 容易にアミノ酸配列にtranslateできるものであれば、どんなものでもよい
- 解析プログラムにかけやすいformatにしてほしい
- 国際統一規格ができればよい
- 早く世界統一のデータベースになってほしい

- GenBank formatでよい
- PCしか使えないので（能力的に）そのような形での配布を希望

DNAデータベースを利用していますか？

YES 72 (57 / 15) NO 34 (32 / 2) 無回答 7 (5 / 2)

YES の場合、DDBJからDNAデータベースの配布を利用していますか？

YES 35 (26 / 9) NO 34 (29 / 5) 無回答 1 (0 / 1)

NO の場合、その理由は？

計算機センター等のDNAデータベースを利用 17 (17 / 0)

直接GenBank. EMBL. NBRFからデータを得ている 10 (6 / 4)

何故ですか？

配布が遅い	0	
その他	DDBJの存在を知らなかった	2
	DDBJへのコンタクト法を知らなかった	1
	DDBJからの配布について最近知った	1
	DDBJの体制が整うのを待っていた。今後は利用する	1

その他

- まだデータが少ないので知人に依頼している
- 解析用ソフトとデータを合わせて購入
- 業者配布のデータベースを利用
- ソフトメーカーを通じて購入している
- 市販のCDのデータベースと、計算機センターのデータベースを併用
- 今後 DDXによる利用が可能になれば利用する予定
- 市販品を利用
- 市販のCD-ROMデータベースを利用 - 検索スピードが遅くデータの更新も遅れるが手軽で簡便

YES の場合、利用状況をお知らせください

- 三井情報のDNA 配列分析ソフトを用いて、パソコンで使用
- 配列データとGenBank floppy版に付録のプログラム、自主開発のプログラムを利用してデータの抽出と解析
- 容量等の点からまだ利用できる状態になっていない
- 学会発表
- データベースから直接配列を引き出して解析に利用
- 良く利用している 3名
- 希望遺伝子の配列ファイルを見る
大型計算機にデータベースを構築することの準備

- FD版を利用しているので便利とはいいがたい。早くDDBJに直接アクセスできるようになってほしい
- IDEAS UWGCG 等を利用したData base 検索に威力を発揮
- データの更新をみている
- 既知の遺伝子とのホモロジー比較、特定Sequenceを引き出しmapping のチェックに使用。また特定遺伝子のクローニングの際の実験計画の資料にも使用
- 週に数回ホモロジー検索に用いている
- 検索、DNA の翻訳
- データベース及び解析ソフトを十分に使いこなせる人の協力のもとに十数名の研究者が利用
- アライメント、サーチ
- 肝炎ウイルス遺伝子配列
- 光合成関連遺伝子の検索、Oncogeneの検索、酵素・生理活性蛋白質遺伝子の検索
- 比較、検索を行っている。当方でかなりextensive な書換えを行った上で利用すべくこれらの目的のプログラムを開発して対応している
- 現在、解析ソフトが良くないので効率良く使われていないが、良い解析ソフトを入手できる予定なので、利用が増えるのは確実
- 金久先生のIDEAS で配列解析を行っている
- 遺伝情報解析の研究において、解析ソフトウェアの開発に利用
- VAX/VMS でIDEAS を使用して利用。GenBank.EMBLの両者とも検索した方が便利
- 基本的なデータベースとして利用。現在解析システムを開発中
- 他プロジェクト（調査研究業務）における資料
- DNA 、あるいは蛋白のホモロジー検索
制限酵素による切断部位
- 金久先生のIDEAS で利用している

磁気テープ配布を希望した営利企業に少額の非強制的な寄付を募ることを計画しています。ご意見をお寄せ下さい。（営利企業の場合のみ、回答）

- 妥当
- データベースによる研究内容は営利企業であろうと、非営利団体であろうと変わらないので、費用が必要なら一律にすべきである
- かまわない
- ”寄付”より”磁気テープ配布料”という形の方がやり易い。GenBankやNBRFと同等の料金を請求する体制にしたらどうか
- 請求書の形の方が払い易い
- できる限り協力したい
- 企業が費用をsupport して配布業務をスムーズに行う事を希望
- やむを得ない 2名
- ニュースレターなど定期刊行物の代金として請求した方が払い易い（磁気テープの配布

は希望していないが、システムの完成を援助したいので)

- 寄付は困難なので有料制を望む
- 発送料等の必要経費を年会費（できれば10万円以内）のような形で徴収するとよい
- 多少の費用負担があっても、恒久的なサポートをすることを希望する
- 寄付という形ではなく、手数料として明確に金額を表示する方が利用する方は金銭処理が煩雑ではないと思う

DDBJの DNAデータベース配布に関するご意見

- フロッピーによる配布を続けてほしい
- 配布のチャージを年2~3万程度集めて、もっと頻度を上げてはどうか
- パソコンで検索しているユーザーのために、NBRFの蛋白データベースについてもフロッピー版を配布してほしい
- Cassette MT (Micro VAX II) での配布を希望
- 電話等でも利用できるようにしてほしい
- 大変有難い制度
- 親切で心のこもったサービスに感謝している
- DDBJのコンピューターに（ユーザーの）インテリジェント端末からACCESSできるようになれば、必要なくなると思う
- データベース利用や遺伝子解析のソフトを充実させてほしい
- compact diskによる配布が早期実現するよう強く希望
- 将来on-line にすることが望ましい
- パソコンレベルで解析できる内容であってほしい
- 日本にはこの様な機関は必要であり、その重要度は一層増すであろう
- 少なくとも送料など、DDBJで負担しているものはuser側に負担させるべき
- DATAのformatをGenBank かEMBLのどちらかに統一、あるいは全てを一つのformatに統一してほしい
- 企業、大学を問わず、開かれた組織でなければならない
- 生物工学解析システムの基礎データベースとして期待している
- 配布時に種々のformatのテープが出力できるシステムを望む
- いつもありがとうございます
- IDEAS 等の検索システムも取り扱ってほしい

DNA データベースに

個人として興味がある	67 (56 / 11)
計算機の管理者として、利用者へ提供する立場からの興味	14 (8 / 6)
無回答	39 (34 / 5)

※複数回答有り

所属機関に利用できる計算機はありますか？

YES 80 (69 / 11) NO 18 (17 / 1)

DDBJ計算機システムの利用を望みますか？

YES 63 (57 / 6) NO 33 (26 / 7)

NOの場合、その理由

DNA データバンクとは直接研究が関係しない	3 (3 / 0)
所属機関に計算機がある	14 (9 / 5)
利用できるソフトウェアが無くDDBJ計算機システムの利用を望むメリットがない	9 (8 / 1)

その他

- 希望はあるが、手続上（課金、ネットワーク加入等）むずかしい
- 現状ではこちらの準備が整っていない
- 所属機関の計算機が回線とつながっていないので、利用できない
- 現時点では必要としない 2名
- 計算機システムについての知識が殆どないので、いちいち他学科の専門家に面倒をみて貰わなければならないので
- 今のところ研究室のパソコンレベルで間に合っている。将来も32ビットパソコンが出れば間に合うと思う
- おそらく通信経費が多額になると思われるので

NOの場合、将来DDBJ計算機システムの利用を望むと思われませんか？

YES 21 (17 / 4) NO 8 (7 / 1)

DDBJ計算機システムの利用希望者へのアンケート

課金システムについてご意見をお寄せ下さい

- 課金自体は問題ないが、諸機関からの多数のアンケートに課金状況（支払い状況）を記入するのが面倒
- テープ等の配布を受ける事になれば、無料というわけにはいかない
- 一般の私企業への支払いに準ずるような方式
- できれば無料サービスを続けてほしい。課金の場合、システムの利用料はできる限り抑えてほしい
- 安いほどよい
- 大学や企業に付属していない研究所は無料とし、企業関係は有料とすべき
- がんセンターでの研究費が使えるようお願いしたい
- やむを得ない

- 使用時間について課金 2名
- なるべく安いものであることと、使用機器等の差がない方が望ましい
- 無料
- 科研費消耗品として扱えるとよい
- できるだけ安くしてほしい。システムとしては、校費から払う方法が良い
- 課金システムでよい 3名
- 通常の大学の大型センターと同じ位のコストならよい
- 基本料金なし。検数ごとに課金
- 従量制にしてください
- なるべくfreeがよい。共同研究という形式で利用
- 校費、科研費、それ以外の研究費のいずれでも支払えるようにしてほしい
- 課題申請時に申請内容によってはDDBJサポートの無料課題の枠をつくる。その他は、大学の大型計算機に準じた課金とする
- 大学の大型計算機センターと遺伝研との間に適当な関係が低コストで可能になることを望んでいる

DDBJシステムの利用方法は

モデムを用い公衆電話回線 24 (23 / 1)

300 baud (ccitt v. 21) ですか? YES 12 (11 / 1) NO 5 (5 / 0)

1200 bps , v. 22

パケット公衆回線 12 (11 / 1)

その他 未定 7名 NTSSでつながりませんか?

所属機関のソフトが十分対応していないために、利用できれば、と個人的に思っている

将来計画としてモデムを用いて

共同研究

N1ネットワークから参入を希望 2名

3Dグラフィック表示システムソフトウェアが充実したら利用したい

可能な場合、UNIXネットワークへの加入を望みますか?

YES 61 (57 / 4) NO 13 (9 / 4)

BITNETネットワークへの加入を望みますか?

YES 24 (57 / 4) NO 37 (33 / 4) 未定、保留 4 (3 / 1)

検討中 1

DDBJ計算機システムの管理運営についてのご意見

- 利用セミナー（初心者向け）や講習会を定期的を開催してほしい
- 大学内研究室としては、大学計算機を通じたNIネットワークでDDBJシステムにアクセス

できれば大変便利

- 使用マニュアルを計算機上にHelpまたはManuとして入力しておいてほしい
- 他のコンピューターシステムとの互換性を持たせて欲しい。(同一ファイルの転送が可能など)
- BITNETの利用価値は大きいですが、24時間運転でないと時差の問題が大きく使いにくくなる
- ニュースレターの発行を頻繁にすると連絡が密になってよい
- 今のままでよい
- できる限り早く、所外利用開始を希望。わかり易いマニュアルを作製してほしい
- 大学間計算機センターのネットワークにつながりませんか？(モデムによる公衆回線は経理担当者がいやがって設置しにくい面がある。)
- ソフトウェアが互いに充実してくれば、open use(有料または共同研究)を望む
- 大変よいシステムになっている。努力は評価できる
- 公開講座、講習会を開くことによって利用の仕方を教授してほしい
センターニュース的な、しかも保存のきく解説マニュアル類の発行
利用者による利用内容の紹介 課題の一覧
- 山口大学ではDDXの回線に乗っているので、そのネットワークによりシステムが使えるとよい
- DNA配列、アミノ酸配列共に歩調を合わせて利用し易いデータベースを作成してほしい

DDBJ の運営に反映させていきたいと思っておりますのでお忙しい中ご面倒と思いますがアンケートにご協力下さいますようお願い申し上げます。下記の宛先までお送り下さい。

宛先： 411 三島市谷田1111、国立遺伝学研究所 遺伝情報センター
 遺伝情報分析研究室 DDBJ係、 Phone： 0559-75-0771 Ext. 647

ふりがな
 氏名 _____ 日付 _____
 ふりがな
 所属 _____ 電話 _____
 ふりがな
 住所 _____

研究分野/業務は DNA データベースに関係する分野ですか？ yes no
 研究分野/業務を以下に簡単にお書き下さい。

[] _____]
 [] _____]
 [] _____]
 [] _____]

yesの場合のみ以下にお答え下さい。

あなたはDNA配列決定のような実験に関係していますか？ yes no

yesの場合

DDBJでは、論文雑誌との協力の下、DNA配列決定の論文は accept された時点で、著者にデータベース登録用のデータ記入をお願いすることを計画しています。提出の媒体は著者にまかせるつもりです。

協力していただけますか？ yes no

No の場合、その理由は？

[] _____]
 [] _____]
 [] _____]
 [] _____]

DNA配列収集についてご意見をお聞かせ下さい。

[] _____]
 [] _____]
 [] _____]
 [] _____]

DNAデータベースの format ならびに annotation についてご要望がありましたらお聞かせ下さい。

[] _____]
 [] _____]
 [] _____]
 [] _____]

DNAデータベースを利用していますか？ yes no

yesの場合

DDBJからDNAデータベースの配布を利用していますか？ yes no

noの場合その理由は？

- 計算機センター等のDNAデータベースを利用
- 直接GenBank, EMBL, NBRFからデータを得ている。

何故ですか？

- 配布が遅い。
- その他

[]]
[]]
[]]

yesの場合、その利用状況をお知らせ下さい。

[]]
[]]
[]]
[]]

営利企業の場合以下にお答え下さい。

現在磁気テープ配布、事務作業等のためアルバイトの女性を雇用していますが、一年を通しての校費による雇用は制度上許されず、他の名目で得た科研費により補っています。この分の費用を捻出するため、磁気テープ配布を希望した営利企業に少額の非強制的な寄付を募ることを計画しています。ご意見をお寄せ下さい。

[]]
[]]
[]]
[]]

DDBJのDNAデータベース配布に関し、ご意見をお聞かせ下さい。

[]]
[]]
[]]
[]]

DNAデータベースに

- 個人として興味があるのでしょうか？
- 計算機の管理者として、利用者へ提供する立場からの興味ですか？

前者の場合以下にお答え下さい。

所属機関に利用できる計算機はありますか？ yes no

DDBJ計算機システムの利用を望みますか？ yes no

No の場合、その理由は？

DNA データバンクとは直接研究が関係しない。

所属機関に計算機がある。

利用できるソフトウェアが無くDDBJ計算機システムの利用を望むメリットがない。

その他

No の場合、

将来DDBJ計算機システムの利用を望むと思われませんか？ yes no

DDBJ計算機システムの利用希望者は以下にお答え下さい。？

課金システムについてご意見をお寄せ下さい。

DDBJ計算機システムの利用方法は

モデムを用い公衆電話回線

300 baud (CCITT V. 21) ですか？ yes no

パケット公衆回線

その他

DDBJ計算機では、ネットワーク機能を重視し、研究者間 mail 用として、公衆電話回線を用いるUNIXネットワークに加入を計画しています。国内の大学、アメリカのほとんどの大学研究所研究者へ mail を発送できます。手紙に比べ、2,3 日で届きます。

可能な場合、UNIX ネットワークを利用しますか？ yes no

ヨーロッパ、アメリカで盛んなネットワークとして、BITNETがあります。BITNETは専用線を用います。そのためヨーロッパ、アメリカの加入大学（ほとんどのmajorな大学）へは 24 時間で mail が届きます。しかし専用線の利用料は一月約20万円で高く、加入は利用者数次第です。

BITNET ネットワーク への加入を望みますか？ yes no

DDBJ計算機システムの管理運営についてご意見お寄せ下さい。

アンケートにご協力いただきまして大変有難うございました。 DDBJ

DDBJ 関連行事日程表

- 1987年 2月 DNA データバンク運営委員会
- 1987年 2月25-28日 EMBL/NIH Workshop "Future Database for Molecular Biology"
(遺伝情報研究センター長 丸山 参加)
- 1987年 3月 1日 共同利用計算機の遺伝研内使用開始
- 1987年 5月 遺伝情報分析研究室 宮沢、GenBank 視察
- 1987年 7月16日 計算機接続用電話外線 5回線を敷設
- 1987年 7月 DDBJ DNAデータ Release 1.0 配布
- 1987年 8月 1日 計算機接続用DDX-パケット 1回線を敷設
- 1987年 8月 共同利用計算機の所外オンライン利用開始
- 1987年 8月 電子郵便開始
(GenBank, EMBLとの間の連絡を電子郵便に切り換える。)
- 1987年11月11日 (所外) DNAデータバンク運営委員会
(所内) DNAデータ研究利用委員会 共同開催
- 1987年11月11-20日 DDBJ-EMBL-GenBank meeting (遺伝情報分析研究室 宮沢 参加)
- 1987年 1月 DDBJ DNAデータ Release 2.0 配布予定
- 1988年 2月 第一回 International Advisory Committee on Biological
Databases ワシントンD.C.にて開催予定

編集後記

データ入力もようやく1.0版を1987年 7月にリリースし、又、共同利用計算機も 8月に所外の方々へのオンライン利用開始のはこびとなり、関係者一同ほっとしています。1988年 1月には 2.0版をリリースする予定です。

米国では、LANL-BBN とNIH のコントラクトが今年で切れ、再審査の結果 LANL-IntelliGenetics が契約を獲得しました。EMBL, GenBankともに、昨年から今年にかけてマネージャーが変わり、新しい動きがはじまっています。国際協力も一段とはづみがつき、データフォーマット、データ構造ともに近いうちに模様替えするでしょう。DDBJは、残念ながら 5年出遅れています。ご支援をお願い致します

1987年11月11日 宮沢三造

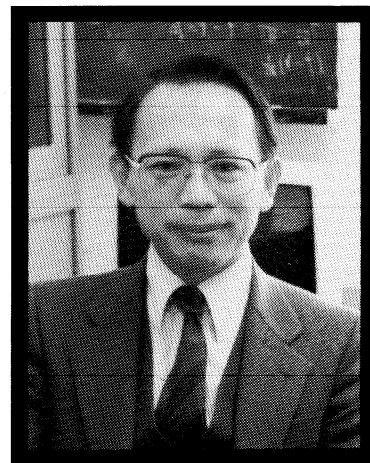
訃報

悲しいお知らせです。遺伝情報研究センター長としてDDBJの責任者であられました丸山教授が、1987年12月11日未明心不全のため 51歳の若さで突然お亡くなりになりました。御遺族の方々に謹んでお悔やみを申し上げますとともに、先生の御冥福をお祈り申し上げます。次ページ以降、先生を悼んで弔辞を掲載させていただきます。何れも 12月14日の告別式に寄せられたものです。

DDBJ一同

丸山毅夫教授の経歴

- 1936年 2月 5日、新潟県に生まれる。
- 1958年 信州大学繊維学部養蚕学科卒
- 1960年 京都大学大学院理学研究科修士課程修了
- 1966年 米国ウィスコンシン大学にて Ph.D.を取得
- 1971年 京都大学より理学博士号を取得
- 1966年～1969年 国立遺伝学研究所研究員
- 1969年～1980年 同研究所集団遺伝学研究系第2研究室室長
- 1980年～1984年 生理遺伝部部長
- 1984年～ 集団遺伝学研究系進化遺伝部門教授、兼遺伝情報研究センター長
- 1987年 12月11日未明、心不全のため逝去



1984年より、我国におけるDNAデータバンク設立に尽力。日本遺伝学会幹事。なお、その間、テキサス大学、ウィスコンシン大学、オハイオ州立大学、ワシントン大学において客員教授を務める。

式辞

丸山毅夫さん、つい四日前まであんなに元気に活躍していた君と、今こうして幽明境を異にして対面することは、まことに痛恨の極みというほかありません。

君は昭和三十三年に信州大学繊維学部を卒業して京都大学大学院に進み、農業研究科修士課程を経て米国に六年間留学し、ウィスコンシン大学で遺伝学の泰斗クロー教授に師事してPh.Dの学位を取得する傍ら、君の持前の数学的才能に磨きをかけるため同大学の大学院数学科修士課程を修めて昭和四十一年に帰国し、国立遺伝学研究所集団遺伝部研究員となり、木村資生部長のもとで集団遺伝学の数理的研究に従事し、とくに地理的に細分化された生物集団の遺伝的構造を支配する法則を解明するなど多くの業績をあげました。また君は、確率過程論に関しては世界的権威の一人で、西独スプリンガー社からこの問題に関する講義ノートが出版されております。こうして昭和四十四年に集団遺伝部第二研究室長に、五十五年に生理遺伝部長に昇任し、さらに五十九年には遺伝学研究所の改組に伴って進化遺伝研究部門教授となり、遺伝情報研究センター長を併任して今日に到ってきました。

君は当研究所に在職の二十一年間、研究者として数多くの勝れた論文を発表して遺伝研の名声を高めてくれましたが、それ以外にもいろいろな面で研究所の運営と発展に貢献してくれました。その第一は、君が着任当初の昭和四十二年ごろ遺伝研に初めて導入されたコンピューター利用の手ほどきに関して所内講習会を木村部長の発案で開催することになり、君は井山審也君や安田徳一君（現在放射線医学総合研究所室長）と共に講師の役を務めてくれました。引き続いて君だけは希望者のために線型代数学を講義することになり、私も含めて数名の有志がこれに参加し、むずかしい演習問題を解くのに苦労した覚えがあります。第二に、毎週水曜日の抄読会では、君はいつも話題の外国論文をわかり易く解説してくれました。生物のパターン形成に関するギーラーの理論や、数学者トムの提唱したカタストロフ理論、生物進化にみられる周期的な種の絶滅と隕石の衝突との関係など、君の抄読会は聞く者にとっていつも教育的で有益でした。第三に、研究所の改組に伴って君は二年間技術課長を併任しましたが、その間技官のために毎週木曜の昼休みに勉強会を主催し、遺伝学の最新的话题を教材として自ら講義をしてくれました。

そして何よりも大きな君の功績は、日本におけるDNAデータバンクの最高責任者としてその創設に献身的な努力を傾けてくれたことでもあります。このデータバンクの設置は全国から要望度が非常に高く、共同利用機関としての遺伝研のいわば使命的な事業になっており、文部省もこれに大きな期待をかけております。

丸山君はこの数年間ほとんどそのために全力投球してきましたが、その努力が実ってようやくこの秋から全国の研究者のためデータベースのオン・ラインサービスができるようになり一同喜んでいた処であります。次の課題としてデータバンクの国際分業をいかに進めるかがありますが、これについての国際会議が明年二月に予定されており、丸山君は日本側代表として出席することになっておりました。その矢先に大黒柱とも頼む君を失ったことは、最愛のご遺族にとっては申すまでもないこと、研究所にとってもかけ代えのない損失であり非常に大きな痛手であります。齢五十一の君にはまだまだ家族のため研究所のため活躍して貰わねばならぬことが沢山あったのです。

雪深い新潟に生まれ育った君は、純朴で誠実な生来の人柄に加えてアメリカ的な善意と親切を信じ、自らもそれを人に施して何ごとにも全力投球して手を抜かず、しかも辛抱強いところがありました。どんなに辛いと思われることがあっても、君の口から不平や不満愚痴の類いを聞いたことはありませんでした。家庭にあっては、君は良き夫であり良き父であると共に最良の家庭教師でもありました。君のこの余りに早すぎる突然の逝去は返すがえすも残念で、今もって信じられないほどであります。しかし君が心血を注いだお陰で動き始めたDNAデータバンクは、継続してさらに発展させてゆかねばなりません。これを利用して恩恵を受ける研究者の数は今後飛躍的にふえるに違いありません。そのことを思うと、創設者としての君の功績はまことに大きいものがあります。その上人々に愛された君の人柄と君から受けた善意と親切とは、多くの人々の心の中に忘れ得ぬ思い出として末長く残ることでしょう。

ここに所員一同に代り君の御霊の安からんことを心から祈り式辞と致します。

昭和六十二年十二月十四日

葬儀委員長

国立遺伝学研究所長

松永 英

弔辞

丸山さんは、私が今からほぼ二十年前に遺伝研に参りました時以来、長い間つきあっていただいていた。その当時、集団遺伝部が二研究室の完全な姿に整い、アメリカでPh.Dをとって間もない丸山さん、今は放医研におられる安田さん、それに私の3人が木村先生のもとに希望に胸をふくらませて集まりました。丸山さんはマジソンのウィスコンシン大学で遺伝学のPh.Dを取得されましたが、もうすぐ数学のPh.Dまでとれるというところまで勉強されたとのことで、数学に関する知識が豊富で、よく食事の後など、雑談の間に数学の話がでて私などとても勉強になりました。その能力が、いかに発揮されたのは、丸山さんが集団遺伝部の部員だった頃なされた集団構造に関する一連のお仕事です。木村先生は、本日はどうしてもぬけられない用事のため残念ながらこの葬儀に参列できないとのことですが、先生によると、丸山さんが京大農学部の大学院生のとき、そこへ集団遺伝学の集中講義にいったさい、熱心で才能ある学生に目をつけたのが丸山さんと木村先生の最初の出会いだったそうです。その後丸山さんをウィスコンシン大学のクロウ教授にPh.D学生として推せんし、さらにその後集団遺伝部の研究員として採用し、数理集団遺伝学の専門家として育てあげられたわけです。

集団遺伝の研究にコンピュータが必要であるということで、木村先生の主導のもとにその申請をしましたが、その時は丸山さんが多く雑務をひきうけて下さいました。はじめ入ったのは東芝の機械で、今から考えますと実に原始的なものです。これが当時の集団遺伝部の研究を進めるのにどれ程貢献したかはかりしれません。ラインプリンターが鳥の鳴くような音を立ててアウトプットを出してきますが、丸山さんとアウトプットが出るのを今か今かと待ちながらラインプリンターをのぞきこんでいた姿がつい昨日のここのように思い出されます。わたしは計算機使用についての知識もほとんど持たずに遺伝研に参りましたので、フォトランに関してはぜひ分色々丸山さんに教えていただきました。その当時は分子進化学も発達しておらず、まだ古典的な遺伝学も充分巾をきかせていた時代で、集団遺伝学も今とはだいぶ様子が違っておりましたが、集団遺伝部の研究は、木村先生がそれ以前からしておられた確率過程に関するものでした。この研究の流れはその後分子集団遺伝学として大きく発展しましたが、丸山さんはこの線にそっ

て拡散モデルの数学理論・特に集団構造についてすぐれた解析をされました。たとえば個体の移住がわずかあれば、中立遺伝子は地域分化を起こさないという結果は、蛋白変異の地域分化を理解するのに大変役に立ちました。その後進化遺伝部の部長となって移られ、更にDNA・データ・バンクの事業をなさるようになり、前のように仕事の議論もできなくなりました。あまりに多忙であられたこと、健康に自信を持ちすぎておられたことからこのような悲しい結果になってしまったことはかえすがえすも残念でなりません。心より御冥福をお祈り致します。

昭和六十二年十二月十四日

集団遺伝部門教授

原 田 朋 子

弔辞

遺伝情報研究センターでDNAデータバンクの運営に関し丸山先生のお手伝いをさせていただいたものの一人として、一言お悔やみを申し上げさせていただきます。

私が丸山先生のお手伝いをするようになりましたのは丁度2年前、先生の御尽力のおかげで遺伝学研究所遺伝情報研究センターにDNAデータバンクが設立された時からです。現在、DNAデータを計算機に蓄積しデータベース化することの重要性を疑う人はいないだろうと思います。しかし、約5年前、ヨーロッパ、米国にDNAデータバンクが設立されたころ、日本では、まだそのような考えは比較的少数派であったように思われます。そんな日本において、丸山先生はいち早くその重要性を認識し、日本にもデータバンクを設立するよう他の関係者とともに国に働きかけておられました。また、DNAデータバンクの運営機関として、共同利用研究所である遺伝学研究所が最もふさわしいとお考えになり、データバンクを運営していけるよう下準備を進めてきたと伺っております。遺伝研にDNAデータバンクがおかれるようになったのもひとえに丸山先生のような献身的な御努力のおかげだろうと思います。

DNAデータバンクが設立後は、丸山先生、遺伝情報研究センターのセンター長として公式にその舵取りをなさってまいりました。そのおかげで、データバンクも少しずつ機能しつつあります。しかし、まだまだ先生の舵取りを必要とすることが山積しております。海外のデータバンクとの協議、国内の関係諸機関との折衝等、私どもでは処理出来ないことばかりです。私ども、こんな状況で先生を失い全く途方にくれる思いです。

しかし、データバンクの責任者として先生の肩にかかる責任は想像以上に重いものがありました。なにより、今やデータバンクは一個人のプロジェクトを越え、日本のそして更に世界において、生命科学の分野におけるビックプロジェクトとなっております。ヨーロッパ、米国のデータバンクもお互いの協力なしでは、データの収集が発表に追い付けない状況です。そんな現状から、データバンクを維持、発展させねばならない責任が、日本に求められております。そのような責任からくるストレスと多忙は想像するに難しくありません。丸山先生は責任感の強いお方でしたので、激務を進んでお引受けになりました。勿論、日頃体力には自身のある先生でしたので、無理をなさったのかもしれませんが。しかし私には責任から来るストレスが今日のような結果になったように思われてなりません。お倒れになる一週間前は、米国のデータバンク関係者を研究所に招待するため京都まで出張し、また北海道大学で開かれた遺伝子実験施設連絡会議に遺伝情報セン

ター長として御出席になり札幌からお戻りになられたばかりでした。京都へは私が出張する予定だったのですが、先生が替わって下さいました。私としてはまことに後悔に耐えない気持です。

先生は、データバンクにとって掛替えのない人でした。まして御専門の集団遺伝学、進化遺伝学の分野ではそれ以上だったと思います。私も幸いなことに、学生の頃丸山先生の集団遺伝学における確率モデル、特にstepping stone modelに関する論文を読む機会に恵まれましたが、その深い洞察に圧倒される思いでした。そんな先生が51の若さでお亡くなりになりました。さぞ無念であったろうと思います。もっともっと、お仕事をさせていただきたかったと切に思います。それだけが残念です。

そんな先生の御遺志は、進化遺伝研究室の五條堀さんを始め、集団遺伝系の方々が立派に後をお継ぎになることだろうと信じております。データバンクに関係する私共は、せめて先生の墓前に喜んでもらえるような報告ができるようデータバンクをもりたてていく以外にありません。どうか先生も我々を見守って下さいますようお願いするばかりです。

つたない言葉で失礼いたしました。これで私の弔辞にかえさせていただきます。

1987年12月14日

遺伝学研究所
遺伝情報研究センター
遺伝情報分析研究室
宮澤三造

以下は、丸山教授の恩師の一人である Crow 教授御夫妻、親友である Felsenstein 教授、集団遺伝学を通じての知人である Cockerham, Langley, Weir 教授、Levin 教授、また GenBank の manager である Burks 博士から寄せられた弔辞です。Crow 教授御夫妻からは電話で、その他は電子郵便で宮澤宛寄せられたものです。PIR の責任者 Barker 博士からの手紙と合わせ掲載いたします。

We join many friends of Takeo Maruyama expressing our deepest sympathy to his family. Takeo's life was cut short at the very highest of his career. First as a student, then as a colleague, he has been a dear friend. We shall miss not only his science but his lively cheerful presence.

Jim and Ann Crow
(Professor, Univ. of Wisconsin, Madison)

FOR FUNERAL OF TAKEO MARUYAMA

I am very saddened to hear of the sudden and untimely death of my good friend Takeo Maruyama. I knew him since about 1962, when I began coming to the laboratory of Jim Crow as an undergraduate student. Takeo Maruyama was then a graduate student in that lab. Over the years I have had frequent contact with him, especially during my visits to Japan at the time of the International Congress of Genetics in Tokyo in 1968, and during the summer I spent in Mishima in 1972, when he and his wife were very helpful to me.

Scientifically I think of him first as the author of the long series of papers on the geographical distribution of neutral mutations (stepping stone models in particular). These papers will long remain essential reading for anyone interested in the population genetics of neutral genetic variation. More recently he had been organizing efforts to make the molecular databases of DNA and protein sequences available on a computer network to scientists in Japan. This was a very exciting development in his career, and it is sad to think that he did not live to see it developed more completely.

I think that his influence on others extended far beyond his scientific work. In particular what made a strong impression on me was his sense of humor and his kindness. He always seemed to have a humorous comment on the life around him. For example, I remember that at his speech at the recent symposium in Madison in honor of Jim Crow, when the slide projector failed, he immediately suggested that the next time they should buy a Japanese projector! He was a keen observer of American life and culture, and he seemed to have an unusually good appreciation of the way American science worked, and which behaviors would be considered appropriate in one system but not in the other.

I remember with particular fondness the times during my visit in 1972, when he and his wife were kind enough to invite me to their house -- at that time one of the small houses owned by the institute -- to use their washing machine to wash my clothes. I would then usually be invited to stay for dinner. They knew that, even though this was not considered usual behavior in Japan, it was in the U.S. It gave me a wonderful picture of Japanese family life which I would not otherwise have had. Although he understood American culture well, Takeo Maruyama always seemed to me to be very Japanese. His wisdom, sense of humor, and kindness have been a wonderful example to me of the strength of Japanese culture.

He was a good and valued friend. We will all miss him very much.

Joseph Felsenstein
Professor
Department of Genetics
University of Washington
Seattle

Dear Colleagues of Takeo Maruyama,

We were saddened to learn of the untimely death of Dr Maruyama. Our field has lost a valued contributor, and Japan has lost a prominent scientist. Please convey our condolences to his family. Our thoughts are with them, and with you, at this time.

Sincerely,
Clark Cockerham, Charles Langley and Bruce Weir

(North Carolina State Univ.; NIEHS; North Carolina State Univ.)

Amherst, December 12, 1987

Dear Mrs. Maruyama:

I was very sad to hear of Takeo's death.

I did not know Takeo well. We only met on two occasions, once in Houston and once in Amherst, when when he stayed with my family and myself while visiting the University of Massachusetts. I liked him very much. I am very sorry that I did not have more opportunities to know him better. I had always thought I would.

I, of course, knew Takeo's research. He was a superb Population Geneticist. His contributions have been important for all of us in this field. Takeo's passing at this early age is a great loss, a tragedy in the true meaning of the word.

I grieve for you and your children and offer you my deepest condolences.

Takeo Maruyama was an excellent Scientist wonderful person. His work will remain important for generations to come. And for all of use who have been fortunate enough to have known Takeo, he will be remembered with great fondness.

Sincerely yours,

Bruce R. Levin
Professor

(Univ. of Massachusetts)

Sanzo,

I am very sorry to hear of Dr. Maruyama's death; it was a pleasure to have met him last year in Heidelberg. We very much appreciated the initiative he tooks in linking DDBJ up with the GenBank and EMBL efforts. I wish that I could have had the opportunity to get to know him better.

- Christian Burks

(GenBank manager at LANL)



PROTEIN IDENTIFICATION RESOURCE

NATIONAL BIOMEDICAL RESEARCH FOUNDATION
Georgetown University Medical Center
3900 Reservoir Road, N.W.
Washington, D.C. 20007

December 30, 1987

Dr. E. Matsunaga, Director
National Institute of Genetics
Mishima
Shizuoka-Ken 411
JAPAN

Dear Dr. Matsunaga:

It caused me great distress to hear of the sudden death of Professor Maruyama. He was very kind to me, making a special trip to Kyoto to escort me to the National Institute of Genetics for my visit on December 4. He seemed so healthy and vigorous that I could hardly believe the sad news. I am glad to have had the opportunity to know him, even though briefly.

I much enjoyed my visit to your Institute. I am glad to have talked also with Dr. Miyazawa. He is very knowledgeable about the problems of databases, and I found his opinions and ideas much in accord with mine. I must say that I also formed the impression while visiting the new GenBank contractor in California that they too hold him in high regard. Thus, I think that the cooperating international databases will have no fears about the future of DDBJ in spite of the tragic loss of Professor Maruyama.

With best regards,

Sincerely yours,

Winona C. Barker, Ph.D.

Mailing Addresses for Inquiries

If you have any inquiry, please send mails to the following addresses (...@niguts.nig.junet).

nig	about the nig system
postmaster	about mails, including address representation.
ddbj	about DDBJ activities
ddbjsub	data submission to DDBJ
genbank	inquiries to GenBank
gbsub	data submission to GenBank
embl	inquiries to EMBL
emblsub	data submission to EMBL

Mails to genbank, gbsub, embl, and emblsub will be forwarded to them.

NOTE: On the nigvms, type "\$ smtp ...@niguts.nig.junet".

DNA Data Bank of Japan 所在地

411 三島市谷田1111
国立遺伝学研究所
遺伝情報研究センター
遺伝情報分析研究室内 DDBJ
電話 0559-75-0771 (内) 647
E-mail: ddbj@niguts.nig.junet

DNA Data Bank of Japan スタッフ

General Manager	遺伝情報研究センター長	丸山毅夫 (tmaruyam)	(E-mail:)
Manager/Database administrator	遺伝情報分析研究室	宮沢三造 (smiyazaw)	
Scientific Reviewer	遺伝情報分析研究室	林田秀宜 (hhayashi)	
関連事務担当	DDBJ 非常勤	堀江元乃 (dbs)	
			(@niguts.nig.junet)

