

ンハイブリダイゼーションを行ったところ、中部および、後部絹糸腺で転写されていることがわかった。

(4) マウスホメオティック遺伝子の発現とクロマチン構造 (浦・廣瀬): マウス第 11 染色体にはホメオボックスを持つ遺伝子 Hox 2.1~2.7 がクラスターをなして存在し、発生における形態形成に関与していると考えられている。Hox 2 遺伝子群の発現調節について調べる目的で以下の研究を行った。マウス F9 細胞をレチノイン酸処理すると、処理前には検出されなかった Hox 2.1 および、Hox 2.3 mRNA が出現した。Hox 2.1 遺伝子の発現は DNA トポイソメラーゼ II の特異的阻害剤 VP-16 により抑えられ、レチノイン酸による発現誘導に DNA の高次構造が関与していることが示唆された。DNA とトポイソメラーゼ II の共有結合中間体を固定することにより、トポイソメラーゼ II の結合部位を探索したところ、Hox 2.1 遺伝子の近傍には検出されなかった。現在、パルスフィールド電気泳動を駆使して Hox 2 クラスター全域にわたってトポイソメラーゼ II 結合部位のマッピングを行っている。

(5) アデノウイルス初期遺伝子の転写制御機構の解析 (半田・廣瀬): アデノウイルス E4 遺伝子の上流域には E4TF1 および、E4TF3 と名付けたタンパクが結合し、エンハンサーとして機能する。これらの因子の c-DNA をクローニングする目的で、精製したタンパクの N 末端アミノ酸配列を調べたが、N 末端がブロックされていて配列を決定できなかった。その他、アデノウイルス E4 遺伝子の転写に関する研究を行った (Handa, H. *et al.* (1989) FEBS Letters 249, 17-20)。

(6) SLE 患者血漿中の抗原 DNA の分子生物学解析 (寺田・廣瀬): DNA を抗原とする SLE 患者血漿から抗体・DNA 複合体を分離し、DNA を抽出してクローニングした。得られたクローンの塩基配列を解析したところ、トランスポゾン Tn3 の  $\beta$ -ラクタマーゼ遺伝子や大腸菌 metK 遺伝子の一部と高いホモロジーをもつクローンが存在した。SLE 患者血漿中の免疫複合体に細菌由来の DNA が存在することは、抗原 DNA の起源について重要な示唆を与えるものと考えられる。

(7) 哺乳動物細胞の DNA トポロジーに関する研究 (岡田・廣瀬): マウス FM3A 細胞から分離した VP-16 耐性変異株および、野生株より DNA トポイソメラーゼ II を精製し、比較した。その結果、超らせん DNA の弛緩活性では両者にはほとんど差がなかったが、catenation 活性については変異株由来の酵素の方が野生株の酵素よりも VP-16 による阻害に対し抵抗性を示した。従って、この耐性株ではトポイソメラーゼ II 遺伝子自身に変異が起きているものと推定される。

#### G-d. 遺伝情報分析研究室

当研究室 (宮沢・林田) は、遺伝情報の研究及び DNA データバンクの運営 (データベース構築, 検索, 解析プログラムの開発, データベースの配布と計算機のオンライン利用のサポート, ニュースレターの出版等) に従事した。

1989 年 2 月 3 日~4 日 EMBL (西独, ハイデルベルク) で第二回データバンクのた

めの国際諮問委員会が開かれ DDBJ を代表して宮沢が参加した。また 6 月 19 日-23 日に各データバンクのスタッフを迎えデータバンクの実務協議が三島の遺伝研で開かれた。議題は関係データベースへの移行スケジュール、データベースの同一性を如何にして保つか、研究者自身によるデータ入力のための支援ソフトウェア、データ収集における地域分担方式への移行等である。またこれを機会に 6 月 24 日 DNA Databases and Genome Projects 研究集会をもった。

(1) 日本 DNA データバンク (DDBJ) 活動

(i) ニュースレターの発行 (宮沢・林田): DNA データバンク活動の報告のため、1989 年 3 月ニュースレター No. 8 を発行した。今年度 466 部配布した。

(ii) DDBJ 計算機利用デモンストレーション: 10 月 23 日-24 日癌学会 (宮沢・林田) で計算機利用によるデータ提出及び解析プログラムの使用に関してデモンストレーションを行った。また 11 月 3 日~6 日生化学学会 (瀬野センター長)、11 月 29 日~12 月 2 日分子生物学会 (林田) ではデータバンク利用のための資料を配布した。

(iii) DNA データベースの導入 (宮沢・林田): 米国から GenBank データベース、欧州から EMBL, SwissProt 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。磁気テープの配布総数は 450 本である。

(iv) DNA データベースの構築 (宮沢・林田): 1989 年 1 月に 4 版 (302 エントリー, 535,985 塩基), 1989 年 7 月に 5 版 (395 エントリー, 679,378 塩基) をリリースした。データ収集の地域分担方式に向け、Genbank 担当の論文雑誌についてもデータの入力処理を 10 月から開始した。今後合意ができ次第 EMBL 担当の論文雑誌にも拡大する計画である。

(v) DNA データ入力、管理システムとデータ検索システムの構築; ネットワークデータベースサーバー機能と電子メールによるデータベースの自動更新機能の追加 (宮沢): データベース検索ソフトウェア FLAT にコマンドを電子メールで送付することによりデータベースを検索することが可能なネットワークデータベースサーバー機能と電子メールによるデータベースの自動更新機能を追加した。それに伴い、The Manual of the Flat Database and Sequence Analysis System for DNA and Proteins, version 1.2 をリリースした。現在 DDBJ が用いているフラットファイルデータベースのための DNA データ入力、管理システム及びデータ検索システム FLAT の概略は“DNA Data Bank of Japan: Present Status and Future Plans”として発表した。

(vi) 関連データベース (HGML) への利用の機会を提供: エール大学の計算機を用い公開されている Human Genome Mapping Library データベースへのアクセスを DDBJ 利用者に提供。

(2) Basigin, a new member of the immunoglobulin superfamily of broad distribution has strong homology with both the immunoglobulin V domain and

---

\* 鹿児島大学医学部

the beta-chain of major histocompatibility complex class II antigen (T. Miyauchi, T. Kanekura, A. Yamaoka, M. Ozawa, S. Miyazawa and T. Muramatsu): マウスにおいて Lotus tetragonolobus agglutinin (LTA) は early embryonic cells へ結合するがその結合部位の蛋白質の一つとして単離された basigin と名付けられた糖蛋白質が Immunoglobulin superfamily の一員であることがアミノ酸配列の類似性から示唆された。蛋白質部分の分子量は約 30,000 と推測され, major histocompatibility class II antigen の beta chain, 及び Immunoglobulin の variable domain と強い類似性を持つ (各々約 120 塩基と 80 塩基の) 配列を含む。また保存されている二つのシステイン残基間の残基数 (62 残基) は Immunoglobulin の variable domain の場合より短く constant domain の場合より長い。免疫グロブリン遺伝子族の進化を考える上で興味深い蛋白質である。

(3) Nucleotide sequences of immunoglobulin-epsilon pseudogene in man and apes and their phylogenetic relationships. (Ueda, S., Watanabe, Y., Saitou, N., Omoto, K., Hayashida, H., Miyata, T., Hisajima, H. and Honjo, T.): ヒト科に属する 3 種のサルとヒトの系統関係を明らかにするため, 免疫グロブリン IgE の heavy chain constant region の偽遺伝子である Epsilon-3 遺伝子の配列の解析を行った。この遺伝子は Processed type の偽遺伝子で, この遺伝子に働く機能的制約は考える必要がなく, 生み出されると同時に機能を失ったことが明らかという点からも系統樹を作成するのに都合の良い遺伝子である。我々の解析の結果, 以前からいろいろな議論のなされている問題であるヒト, ゴリラ, チンパンジーの系統関係は従来からの分子進化学的知見と一致し, ゴリラがヒトとチンパンジーの共通の祖先から分かれた後ヒトとチンパンジーが分岐したという結果を得た。この結果は統計的には有意ではないが, 従来からの分子進化学的知見を補強する一つの新しいデータである。

### G-e. 遺伝子ライブラリー研究室

3 月より, 小原雄治助教授が担当教官として着任し, 永田妙子が研究補助業務をおこなった。

本研究室では, 遺伝子ライブラリーの構築, 管理, 配布という業務と, このための新しい方法論の開発を行い, 並行して, 遺伝子ライブラリーを活用して動物発生過程の遺伝子発現ネットワークの解明をめざす研究を始めている。

(1) 大腸菌遺伝子ライブラリー (小原・永田): 小原が名古屋大学在職中に作成した大腸菌ゲノムの遺伝子ライブラリーの維持, 配布, 情報収集及びそのデータベース化を行っている。このライブラリーの特色は, 個々のクローンについて詳細な制限酵素地図が作成されており, これをもとに, 大腸菌全ゲノム 4700 キロ塩基対が, 互いに少しずつオーバーラップするクローンでおおわれていることである。遺伝地図との対応づけができていたので, ゲノム上のあらゆる場所へのアクセスが非常に容易になった。1~数千クローンを調べればよいからである。本研究室では, 総数 3400 クローンの中から十分な重なりをも

- (4) 広田の大腸菌温度感受性変異株のコレクション: 約 5,000 株
- |              |           |
|--------------|-----------|
| DNA 複製欠損変異株  | 115 株     |
| RNA 合成欠損変異株  | 100 株     |
| ムレイン生合成欠損変異株 | 55 株      |
| 細胞分裂欠損変異株**  | 353 株     |
| 染色体分配欠損変異株** | 45 株      |
| 膜蛋白欠損変異株     | 22 株      |
| リボソーム蛋白変異株   | 79 株      |
| 未同定欠損変異株     | 約 3,800 株 |

\*\* "Mapping of a whole set of cell division genes in *Escherichia coli* K-12" (1991 年 2 月に出版予定)

- (5) *Escherichia* のファージ: T<sub>2</sub>, T<sub>3</sub>, T<sub>4</sub>, T<sub>4</sub>GT7, T<sub>5</sub>, T<sub>6</sub>, T<sub>7</sub>, Pl·kc, Pl·vir, Mu, λpapa, λvir, λgt·λC, λcb2, λcI<sub>857</sub>·S7, λTn5, λTn10, φX174wild, φX174am3, f1, MS2, Qβ, その他

### その他

*Bacillus subtilis* (枯草菌): 200 株

*Salmonella typhimurium* (ネズミチフス菌): 1,370 株

## II. 遺伝情報の収集保存

現在 DDBJ で利用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。

### DNA 塩基配列データ:

DDBJ	5 版 (07/89)	395 エントリー	679,378 塩基
EMBL	21 版 (11/89)	28,679 エントリー	34,748,087 塩基
GenBank	62.0 版 (12/89)	31,228 遺伝子	37,183,950 塩基
NBRF	35.0 版 (05/89)	3,205 エントリー	7,151,795 塩基
HIV-N	1988 年版		
KABAT	1983 年版		
Miyata	1988 年 3 月版		

### 蛋白質アミノ酸配列データ:

DDBJ	5 版 (07/89)		
PIR	23.0 版 (12/89)	6,550 蛋白質	1,942,966 残基
SWISSPROT	12 版 (10/89)	12,305 蛋白質	3,797,482 残基
KABAT	1983 年版		

コドン使用頻度データベース: (GenBank 50.0 版に対応)

LiMB (Listing of Molecular Biology Database):

1 版 (02/88)

以下は各データベースの簡単な収集内容である。

### 1. GenBank Release 62.0

グループ	エントリー数	塩基数
盤 長 類	5,340	6,412,545
ゲ ッ 菌 類	5,293	5,400,772
哺 乳 類	1,061	1,329,460
脊 椎 動 物	1,431	1,616,292
無 脊 椎 動 物	2,264	2,827,148
植 物	2,133	3,178,773
オ ル ガ ネ ラ	960	1,487,642
バ ク テ リ ア	3,015	4,783,028
R N A	1,242	224,740
ウ イ ル ス	2,830	4,743,041
フ ァ ー ジ	476	578,431
人 工 合 成	848	303,520
無 注 釈	4,335	4,298,558

### 2. EMBL Release 21

グループ	エントリー数	塩基数
人 工 合 成	731	262,477
ク ロ ロ プ ラ ス ト	388	853,519
Genetic elements	102	108,762
ミトコンドリア遺伝子群	720	809,608
原 核 生 物	3,353	4,632,129
ウ イ ル ス / フ ァ ー ジ	3,161	4,861,022
真 核 生 物	17,850	20,726,635
そ の 他	59	90,535
無 注 釈	2,315	2,403,400

### 3. PIR Release 23.0

グループ	エントリー数	残基数
真 核 生 物	3,389	781,966
哺 乳 動 物	1,870	458,898
植 物	408	83,835
真 菌 類	212	72,783
原 核 生 物	1,170	303,350
動 物 ウ イ ル ス	1,332	652,958
植 物 ウ イ ル ス	70	30,687
バ ク テ リ オ フ ァ ー ジ	372	83,312