

第4章

生物科学研究の100年

今日ライフサイエンス、生命科学、生物科学などと呼ばれる研究分野は、過去100年の間に名実ともに大きく変容した。理化学研究所における研究もまた、時代とともに大きく変化してきた。その源流をあげれば、有機化学の鈴木梅太郎の研究室に行き着く。

1922（大正11）年の職員名簿には、次のような9項目の研究課題が鈴木（梅）研のところに記載されている。駆虫剤、脂肪酸塩によるアルデヒドの生成、アミノ酸合成、日本酒、脂肪の栄養価、糖蜜利用、内燃機関用燃料（大河内研と共同）、ビタミン、アミノ酸の電解。ここから分かるように、100年前の理研では、まず広い意味での「有機分野の科学」が展開されたのであった。同研究室では、中原和郎研究員による腫瘍（がん）の研究も早い時期から始まっている。

農薬の研究の歴史も古い。1921年には、玄米の害虫コクゾウムシを防除するクロールピクリンが工業化された。戦後の1964（昭和39）年には、農薬用抗生物質のポリオキシンが理研で発見され、現在も安全性の高い主要農薬の一つとして広く使われている。21世紀初頭の現在、理研は日本における生命科学研究の中心とまで言われるようになったが、そこに至る大きな源流は、実は、1962年から約10年間展開された新規農薬の創製プロジェクト（正式には1987年まで）である。これは、その後に登場するプロジェクト研究の先駆であった。

期を画したのは1971年の科学技術会議による第5号答申であった。今日的なライフサイエンスの振興策が叫ばれ、その実施場所が理研となったのである。さらに第8号答申を受けて、1984年に理研のライフサイエンス筑波研究センターが誕生した。この流れの中から、1997（平成9）年の脳科学総合研究センターを皮切りに、多くの研究センターが誕生していった。

その一つが1998年に設立されたゲノム科学総合研究センターであった。2008年3月に解散するまで、いくつもの画期的な成果を上げるとともに、その「ゲノム科学」が、その後の生命科学にまったく新しい研究スタイルを持ち込んだ。生命科学の方法論が明らかに変わったのである。

ここでは理研における生物科学の概史を記した後、この章の補遺として理研のライフサイエンス研究の成功例のFANTOMを取り上げ、いかなる経緯で世界から高い信頼を勝ち得たのか紹介する。

第1節 理研の生物科学の源流

鈴木梅太郎研究室

理研が創設された1917（大正6）年、当時43歳だった鈴木梅太郎は、オリザ



鈴木梅太郎

ニン（ビタミンB1）の研究を中止していた。第一次世界大戦の影響でドイツからの薬品輸入が途絶え、研究を中止せざるを得なくなったからである。

それより前、鈴木は1901（明治34）年から6年間、エステル合成法で知られるベルリン大学のフィッシャー（Hermann Emil Fischer）のもとに留学していた。フィッシャーは1902年にノーベル賞を受賞した後、アミノ酸からタンパク質を作る研究を進めており、鈴木はそこで難しいトリペプチドの合成に成功、大きな貢献を果たしていた。32歳で帰国してから4年後の1910年、駒場の東京大学農学部で、オリザニン（ぬか）を米糠から純粋に取り出すことに成功した。

このころ、3大栄養素以外の必須な栄養物質としてビタミンという概念（と物質）が登場したが、日本でも海外でも、否定する意見が医学界に強かった。しかし、大戦が終了して鈴木が研究を再開した1918年には、京大や慶応の医学部でオリザニンが脚気の治癒に非常に効果的であることが認められ、状況は一変していた。その後、1927（昭和2）年に理研でオリザニンの結晶化に成功した。しかし、構造決定に関しては、アメリカの研究者による1937年の仕事に先を越された。

鈴木梅太郎は理研で最大規模の研究室を運営し、さまざまな研究や事業に関わっていく。以下にその一端を紹介する。

ビタミンAの精製

1922年、鈴木門下の高橋克己がタラの肝油から脂溶性ビタミンであるビタミンAの精製に初めて成功し、「ビオステリン」と命名した。ちなみにオリザニン（ビタミンB1）は水溶性ビタミン。脂溶性ビタミンの単離は世界中で試みられていたがうまくいかなかった。そのような中で、高橋らは巧妙な精製法を編み出すとともに、詳細な動物実験によって、その医療効果や過剰摂取による副作用まで発見し、60ページに及ぶ英語の長論文を発表した。この1925年の論文の共著者に梅太郎の名前はなく、論文で提案した化学式は間違っていた。

高橋は1923年に正式に理研の研究員となり、このビタミンAの業績で1924年に帝国学士院賞を梅太郎とともに受賞したが、腸チフスが悪化して翌1925年の2月8日に死去している。論文の公開もその死と相前後したものとなった。

高橋によるビオステリン（ビタミンA）の抽出法に関して、理研は欧米各国の特許も取得した。国内では「理研ビタミン」の商品名で直接、製造販売して大成功を収めた。これにより、財団理研の初期の財政状況は改善され、組織整備も進められたのであった（第1編第2部第6章を参照）。**88年史**によれば、「ビタミンAは理研の経営面でも貢献し、年間研究予算の半分に当たる約30万円を稼ぎ出した。発明者の高橋には1922年から1930年の間に48万円余の報酬が支払われた。アドソールなどと比べると、この報酬は破格の額である」とある。高橋は1925年に没しており、大半は遺族（妻と子女2人）に贈られたのである。

その後も鈴木（梅）研ではB2、B6、C、D…と、ビタミンの幅広い研究が展開された。製造販売に関しては、やがて1938年設立の理研栄養薬品へ移管され、さらに戦後の混乱を超えて1949年に理研ビタミン（株）の前身が設立された。この会社は2017年現在、東証1部上場会社として経営されている。

スレオニンの発見

鈴木（梅）研の研究テーマの一つに、タンパク質の代用としてアミノ酸混合物を使うという研究があった。しかし、混合物でマウスを育てても、すぐに死んでしまっていた。つまり、アミノ酸混合物では、タンパク質の栄養価の代用はできなかったのである。

そんな中、研究室の前田司朗が1933年、馬肉を硫酸で加水分解した生成物と、水酸化バリウムで加水分解した生成物とを混ぜて飼育すると、マウスが正常に育つことを明らかにした。この研究にはイリノイ大学のローズ（W. C. Rose）という競争相手がいて、ほぼ同時に同様の成果が得られた。さらに3年後、最後の必須アミノ酸であるスレオニンについても、両者はほぼ同時に発見した。つまり、前田は既存のアミノ酸混合物にスレオニンを添加すればマウスが正常に育つことを示した。

この発見は、のちに、アミノ酸点滴栄養にも発展し、手術後に食事の摂れない患者にとって福音となった。前田は高橋克己と同じ33歳で逝った。

理研酒

合成酒の開発のきっかけは、1918年に起こった米騒動で、鈴木梅太郎は、食糧としての米を使わずに日本酒を作る可能性を加藤正二らと追求し始めた。まず、米のタンパク質がアミノ酸に分解され、それが日本酒の味と香りを生み出しているという図式に基づき、米に近い絹糸屑とでんぷんのアルコール発酵を試みた。これが失敗したため、純合成法つまり、日本酒を徹底的に成分分析して、それらの成分を混合して合成酒とする方針へと転換した。

1922年に製造法の特許を取得し、翌年に大和醸造から合成酒「新進」が発売された。その後、1924年に下瀬林太と藪田貞治郎が必須成分のコハク酸の製造に成功した。これも含め、理研酒でエタノールに混ぜた成分は、有機酸、アミノ酸、糖類、無機塩類、そして香料など60種類以上にも及んだ。

1928年には、30万円を投資して、理研内に年産2万石（1万石=1800kl）のモデル工場を完成させた。これとは別に、台中、台北の専売工場で年産5万石が1945年まで生産され、毎年10数万円が鈴木（梅）研の研究費補助となった。1943年時点で、理研の合成酒は、47社、年間生産量76万4000石に及んだ。戦後の昭和30年代もほぼ同じ量が生産された。なお、理研酒工業は1955年に協和発酵工業に吸収され、2002年にその酒類事業がアサヒビールに譲渡された。現在も「利久」はアサヒビールから販売されている。一方、「新進」の大和醸造はメルシャンに買収され、別の名前で合成酒が販売されている。

農薬・天然物化学

冒頭でも触れた貯蔵米に被害を与えるコクゾウムシの問題は、政府から鈴木（梅）研に研究が委託された。研究室の山本亮は、1917年に東大から大原農業研究所に異動し、岡山県特産の除虫菊花の殺虫成分を研究し始めた。そこで発見したピレトリンは、現在も蚊取り線香の殺虫剤などとして世界中で使われている。

コクゾウムシに関しては、第一次世界大戦で毒ガスに使われたクロルピクリンが効くとの発表がフランスであり、鈴木は山本に化学合成を指示した。完成には2年かかったが、当時の三共（第一三共）と一緒に合成農薬を作り、燻蒸試験にも合格し、期待に応えることができた。

その後、山本はピレトリンの研究に戻ったが、化学構造の研究ではスイスのグループの後塵を拝した。それもあって山本はヨーロッパに留学し、帰国後、台北帝大教授として熱帯農産物の研究を深め、そのあと理研に復帰した。戦後は、室員の松井正直とともに住友化学に転職。合成ピレトリンであるアレスリンの合成を完成させ、工業化に成功した。

同じころ、鈴木（梅）研の武居三吉は熱帯植物由来の駆虫剤を研究していた。1922年、デリス（植物）の根や茎から活性成分ロテノンの抽出に成功した。ドイツ留学後に京大教授に就任、1933年にロテノンの構造解析に成功した。

第2節 梅太郎の死と戦後

鈴木梅太郎は1943年に69歳で没した。さまざまな成果を生み、産業化という面でも大きな足跡を残した巨人であった。その後、鈴木（梅）研究室は、鈴木文助研究室と藪田貞治郎研究室に二分されるが、鈴木文助主任研究員は5年後に亡くなった。藪田研究室は、植物ホルモン、農産物の利用研究、合成酒なども引き継ぎ、その活動は多岐にわたって発展していった。

藪田の業績を上げると、ジベレリンの結晶化と構造決定がある。ジベレリンはもともと、台湾総督府農事試験場の黒沢英一技師が世界で初めて発見した植物ホルモンであった。この物質は、種子の発芽促進や老化の抑制など、さまざまなところに関わっている。藪田らの研究以降、ジベレリンの研究は、東大の住木研究室、理研の農薬部門、国際フロンティアなどで活発に進められた。そして横浜研究所の植物科学研究センター（PSC）で、神谷勇治GDらがさらに大きく展開させていった。

ペニシリンの開発

太平洋戦争末期の1944年、稲垣克彦軍医少佐が幹事となって、陸軍ペニシリン委員会という組織が立ち上がった。ここには、理研から藪田貞治郎と久保秀雄（仁科研）が参加、それにのちに理研再興に関わる坂口謹一郎、住木諭介（藪田研〈兼務先の東大での〉の助教授）のほか、後の抗生物質開発で名をあげる梅沢浜夫らも参集した。坂口の回想が残されている。「……戦争末期から終戦直後に

かけての抗生物質、主としてペニシリンの研究のごときは、その後の日本の発酵学会の大きな転機ともなった。藪田先生と住木博士が分離・精製、朝井勇宣博士と私が菌株・培養を担当。幸い、1951年、有馬啓、小笠原長宏博士による色素欠損株の創出などの業績が出た」。

1948年3月、株式会社科学研究所（仁科芳雄社長、藪田貞治郎製造部長）として理研は再出発した。生きていく術の一つがペニシリンの生産であった。理研酒関連施設を活用して生産体制を確立、同年の年間生産量は全国1位となった。ところが過剰生産のために会社が倒産、結核薬ストレプトマイシンの工業化に向けて、生産研究を開始することになった。翌1949年には厚生省のストマイ研究協議会が発足、藪田、坂口、大山義年の各主任研究員が参加した。1951年には北区十条の新工場を完成させ、60tのタンクを整備した。その最中に仁科社長が逝去したのである。

培養は開始したが財政は立ち直らず、1952年、阪谷希一社長は生産部門を切り離して科研化学株式会社をスタートさせると、こちらの方はやがて好転していった。一方、残された科学研究所の状況は厳しく、1958年の特殊法人化まで苦境は続いたのである。

プロジェクト研究の先駆けとなった農薬研究

1959年、日本学術会議は政府に対して、農薬研究の必要性を強調する勧告を行った。当時の日本は食糧不足に悩んでおり、増産と安定した自給体制を目指していた。それには重要作物の品種改良とともに、生産性を高める農薬の重要性、特に新しい農薬が不可欠と考えられた。しかしその研究体制は脆弱だったのである。

この勧告に基づいて、政府は関係各省庁で話し合いをもち、農薬活性物質の探求を中心とする新農薬創製は科学技術庁、農薬の応用・利用に関する問題は農林省、研究・教育の問題は文部省と、対応の大枠が固められた（省庁名は当時）。この結果、科学技術庁傘下にあった理研において、新農薬創製研究体制の構築が試みられることになった。

当時の住木諭介副理事長を中心に議論が進められ、主任研究員制度とは予算枠の異なる農薬研究部門を立ち上げることになった。構想は10研究室に112名の研究員というものだったが、予算折衝の結果、1962年度に定員3名の農薬第一研究室が発足し、毎年一つずつ研究室を新設していくことになった。こうして1970年時点で九つの研究室が置かれた（**88年史**207-216ページ参照）。

- 昆虫薬理研究室（農薬第一研究室）
- 微生物薬理研究室（農薬第四研究室）
- 植物薬理研究室（農薬第三研究室）
- 動物薬理研究室（動物生理研究室）
- 農薬合成第一研究室（農薬第二研究室）
- 農薬合成第二研究室

農薬合成第三研究室
生物試験室
農薬試製室（後の植物化学研究室）

多くの成果を上げた農薬研究であったが、社会的要請が変化したことなどもあり、1987年に解体されることになった。つまり、農薬研究というプロジェクトを構成してきた研究室は、一般の主任研究員研究室と同じ立場で自主的研究を展開することになった。それに伴い、研究室の名称も、昆虫生態制御研究室のように、内容を明示するかたちに変えたのである。

この初めてのプロジェクト研究体制は、既存の主任研究員制度、主任研究員会議とは相容れない部分もあり、さまざまな波紋を広げたという。後から見ると、そのぶん、次なる太陽エネルギー、ライフサイエンス、国際フロンティアといった大型プロジェクトを進める上で、いわば環境整備（地ならし）の役割を務めることにもなった。

第3節 ライフサイエンスの時代へ

六つのプロジェクトの推進

1971年の科学技術会議による第5号答申を受けて、政府はライフサイエンス構想を立ち上げた。しかし諸状況が許さず、法人の設立は見送りとなった。その代わりに、理研の中で全国の研究者を対象とした六つのプロジェクトが推進されることになった。その6テーマは、①老化、②人工臓器、③バイオリクター、④知能機械、⑤生物活性物質、⑥新微生物利用技術であった。

このプロジェクトこそ、今日的な意味での日本のライフサイエンス研究の源流となったものであり、研究レベルの底上げと発展に大きく貢献した。さらに、第8号答申を受けて、1984年に理研のライフサイエンス筑波研究センターが誕生した。

1980年代を迎え、ヒトの遺伝暗号を全て読み解く「ヒトゲノム計画」が持ち上がった。ゲノム科学は、その後の生命科学にまったく新しい研究スタイルを持ち込んだ。理研は、1980年代より全自動シーケンサーの研究開発を進めてきており、その流れに沿って次の展開を模索した。そして、理研におけるゲノムを中心とした研究活動を担うため、1992年に林崎良英が理研ライフサイエンス筑波研究センター・ジーンバンク室に着任した。主任研究員会議メンバーからなるゲノム科学推進グループが組織され、理研における中核的課題として、トランスクリプトーム（全メッセンジャーRNA）からスタートする研究戦略が策定された。

1998年には和田昭允を所長とするゲノム科学総合研究センター（GSC）が発足し、林崎良英、横山茂之、榑佳之の三つのグループで研究をスタートさせた。そして、榑のヒトゲノム、横山のタンパク3000、林崎のFANTOMと、いずれも確固たる足跡と成果を残したのであった。

時代が要請するセンターの設立と変遷

GSCの設立と前後して、理研をライフサイエンス研究の中心拠点とするべく、さまざまな研究センターが次々と設立されていった。医学を含む生物関連分野に絞ってセンターを設立順に書き出すと次のようになる。これには、世紀の変わり目を期したミレニアム研究センター群（下記の*印）も含まれる。

脳科学総合研究センター	1997年10月
ゲノム科学総合研究センター	1998年10月
植物科学研究センター*	2000年4月
遺伝子多型研究センター*	2000年4月
発生・再生科学総合研究センター*	2000年4月
バイオリソースセンター*	2001年1月
構造プロテオミクス研究推進本部	2001年4月
免疫・アレルギー科学総合研究センター	2001年7月
感染症研究ネットワーク支援センター	2005年7月
分子イメージング研究プログラム	2005年9月

その後、さまざまな要因もあってセンターの改廃や組織改編が進められた。中でも大きかったのは2008年の第2期中期計画のスタートに伴うもので、ゲノム科学総合研究センター（GSC）はその役割を終えたとして解体された。そして研究者は、オミックス基盤研究領域、生命分子システム基盤研究領域、生命情報基盤研究部門などへと分散していった。なお、この時の組織改革で、フロンティア研究システムと中央研究所が基幹研究所に統合され、また遺伝子多型研究センターはゲノム医科学研究センターに改称された。2008年10月から分子イメージング科学研究センター、2011年4月からは生命システム研究センターとHPCI計算生命科学推進プログラムが、それぞれスタートしている。

こうした設立・分散の方向から、なかば強引に統合する方向へと組織改革されたのが、2013年の第3期中期計画のスタートに伴うものであった。その統合・組織化は非常に複雑であるが、おおまかに言えば、オミックス・分子構造解析・分子イメージングがライフサイエンス技術基盤センター（CLST）に統合された。また医学系センターは統合生命医科学研究センター（IMS）に統合された。植物科学を含む組織は化学やケミカルバイオロジーなどの分野と統合し、環境資源科学研究センター（CSRS）となった。

2017年現在、理研における生物やライフサイエンス系の研究は、これら3センターに、生命システム研究センター（QBiC）、多細胞システム形成研究センター（CDB）、脳科学総合研究センター（BSI）、バイオリソースセンター（BRC）を合わせた七つのセンターと主任研究員研究室群を中心に進められている。そこでの成果は、第II編で必要かつ十分に記載されている。

以下、本章の補遺としてライフサイエンス研究に多大な貢献をし続けるFANTOMプロジェクトについて、その歴史と歩みを紹介する。FANTOMもま

た、理研における広い意味での生物科学の伝統の中から生まれた独創的成果であり、RNA大陸すなわち、タンパク質をコードしないRNAが多数存在することを発見してセントラルドグマを覆し、FANTOMデータベースを通じて山中伸弥博士のiPS細胞誕生に貢献したのである。

FANTOMプロジェクト

《国際コンソーシアムの成功例》

FANTOMは理化学研究所のライフサイエンス研究の成功例である。FANTOMはいかなる経緯で世界から高い信頼を勝ち得たのか。FANTOM5で20カ国以上、約60機関から多数の人々が参加する世界最大のコンソーシアム（連合組織）となっており、過去22年間、理研および日本が常にリード・運営してきた。

FANTOMデータベースの基本は、マウスのゲノムとRNAに関する百科事典であり、研究者はそれらを参照しつつ研究を進めることができる。タンパク質をコードしないRNA（非タンパクコードRNA；non-coding RNA；ncRNA）の発見でセントラルドグマを覆し、RNA大陸の存在を示したのもFANTOMであった。また、山中伸弥はこのFANTOMデータベースを利用してiPS細胞を発明したと語っている。

第1節 RNA解読を目指す革新的プロジェクト

マウス完全長cDNAプロジェクト

理研で林崎良英のゲノム科学研究所がスタートした1995（平成7）年、世界では、6カ国共同の国際ヒトゲノム配列決定コンソーシアム（HUGO）と、ヒトゲノム配列をビジネスにつなげようとしたセセラ社とが、激しく競争しながらヒトゲノム計画を推進していた。両者は相次いで8年以内にヒトゲノムのドラフトシーケンスを完了する、と発表していた。日本からは、東京大学医科学研究所の榎佳之らがHUGOに参加し、国際協力の立場でヒトゲノムシーケンスを推進する立場を取っていた。そのような状況も考慮しつつ、理研は推進すべきゲノム科学の方向性を決めなければならなかった。

1995年の米国のヒトゲノムシーケンスの予想発表を受けて、林崎は、欧米主導で推進されているゲノムプロジェクトに競合する形の研究を避け、ゲノムに書かれているさまざまな遺伝情報がいろいろな形で読み取られるRNAの配列を解読する、という革新的なプロジェクトを提案した。それには次のようないくつかの理由があったが、それはまた、常にオリジナルな研究を求める理研精神の発露でもあった。

- (1) 欧米が推進するヒトゲノムプロジェクトと競合したり、また、単なる協力者としてヒトゲノムプロジェクトに参加して独自の立場を築けなくなったりするよりは、ヒトゲノムプロジェクトと相乗的かつ補完的効果のあるデータを出すことを考える方が賢明であること。
- (2) ゲノムから転写されスプライシングを受け、さらに転写後修飾（RNAエディティング等）を受けるRNAを読み解かない限り、ゲノム配列だけで

はゲノムに書かれている遺伝情報がどんな形で機能を発現しているのか、知ることができないこと。ゲノムとRNAの配列を比較して、これらの情報が読み解けるという意味で、RNA配列データはゲノム配列データと補完的であること。

(3)完全長cDNAを合成する完璧な技術がなく（当時）、これを開発することで技術的な優位性を保てること。

(4)でき上がった完全長cDNAバンクとデータベースを活用して、さまざまなタンパク質を人工的に発現させてタンパク質構造研究の材料としたり、個々の生命現象の分子メカニズムを追求するタンパク質機能研究の材料としたりすることができ、ライフサイエンス研究の非常に強力なツールになること。

さらに、こうしたトランスクリプトーム解析の材料として、あえてヒトではなくマウスを選んだことにも理由があった。①マウスは、突然変異体も数多くそろい、さらに遺伝的背景が均一な純系が整備されていて、ヒトのモデルとなる実験動物として非常に多くの知見が得られていること。さらに、②体を構成する全ての組織・臓器を用意できること、③倫理的な手続きも不要であることなどがその主な理由であった。このような背景を考え、生体内のRNAを全て読み解こうというマウスゲノムエンサイクロペディア・プロジェクト（完全長cDNAプロジェクト）を提案したのである。

マウスを選んだもう一つの大きな理由は、献身的に林崎をサポートし、この理研プロジェクトに共感してくれたV・チャップマン（Verne Chapman）の存在があった。彼は、国際マウスゲノム学会の創始者であり、アメリカのバッファローにあるロズウェルパークがん研究所の教授だった。突然変異体やゲノム地図の整備状況を考え、C57BL/6マウス純系種を強く推薦してくれたのである。その後、このC57BL/6種が遺伝子配列データベース（完全長cDNA）の基本標準種となった。

チャップマンは共同研究のために何度も理研を訪れたが、理研滞在中の1995年夏に、宿舎内で夜間突然の発作に見舞われ急逝された。チャップマンの弟のG・チャップマン（Gary Chapman）は、アメリカから遺体の火葬許可証を持ち、理研で執り行った葬儀に出席、遺骨を故郷に持ち帰ってくれた。その後、G・チャップマンは、理研ゲノムプロジェクトのクローン技術を用いて設立された理研ベンチャーである（株）ダナフォームの初代CEOとなる。

林崎はまず、完全長cDNAのクローンセットを用意し、RNA配列を読み解くために必要な技術開発を急いだ。当時イタリアからのポストドクであり、後のライフサイエンス技術基盤研究センター機能性ゲノム解析部門長となるカルニンチ（Piero Carninci）とともに、一連の完全長cDNA合成技術を開発した。これらの完全長cDNAライブラリーの合成技術の中には、①完全長cDNA伸長法（逆転写酵素が第一鎖cDNAを完全長に伸長させる技術）、②完全長cDNA選択法（それでも生じた不完全な第一鎖cDNAを除去する技術）、③完全長cDNA正規化法・差分法（各種遺伝子配列をコードするcDNAの出現頻度を均一化して、

同じcDNAばかりを単離しないようにする方法)、④長鎖cDNAクローニングベクター、がある(時計を先に進めると、完全長cDNA技術は、PCR法やサンガーシーケンス法などとともに2007年に*Nature*誌が発表した22個の「DNA技術の金字塔」の一つに選出された)。

1997年当時、新型シーケンサー開発においては、全てアメリカ企業が主導して開発がなされていたが、理研は新型シーケンサーが販売されるのを待たず、独自の大容量シーケンスシステムRISA (RIKEN Integrated Sequence Analyzer)を開発していた(『88年史』331ページ参照)。

このシステムには、当時の研究基盤技術部と共同で開発した1日で4万個のプラスミドを大腸菌から全自動で精製するシステム(RISA全自動プラスミド抽出機)や、島津製作所と共同開発したRISA384多本架キャピラリーシーケンサーが含まれる。これらのRISAシステムは次世代シーケンサーが登場するまで、完全長cDNA解析などの大規模シーケンスを支えてきた。なお、科学史の記録として、現在これらの機器は国立科学博物館(上野)に寄贈されている。

これらのシステムを全て統合し、河合純(現在、予防医療・診断技術開発プログラム副プログラムディレクター)、島津製作所より移籍した秋山純一(上級研究員)が中心となり、理研筑波ライフサイエンスセンターにRISAシステムを使ったシーケンスセンターの建設、運営、運転が始まり、実際のマウスのさまざまな組織の完全長cDNAシーケンスデータが得られるようになった。

アメリカで発表すると、分与・供与の依頼が殺到

完全長cDNAプロジェクトは、1995年から第一目標達成まで、着実にデータを積み上げていったが、その成果は何年間も未発表のままであった。

1999年、理研の産業連携担当として着任したG・チャップマンとともに、林崎は、アメリカのコールドスプリングハーバー研究所(CSH)で毎年開かれている「ゲノム・シーケンシング」というタイトルのシンポジウムに出席した。林崎の発表予定はなかったが、友人で当時ニューヨーク大学(NYU)にいたシュワルツ(David Schwartz)が、今まで理研が進めてきたマウスゲノムエンサイクロペディアプロジェクトについて、ぜひ飛び入りで話をするよう強く推薦し、組織委員長と2回にわたって掛け合ってくれた。

しかし、聞き入れられずにいたところ、あるセッションの発表者7人が一人1分ずつ短くしてくれる交渉をシュワルツとG・チャップマンが取り付け、晴れて7分間の発表の機会が与えられることになった。その結果、林崎は、それまでの4年間に開発した完全長cDNAシステムについて、また目いっぱい溜め込んできた大量の知見について、直前1時間の準備だけで、7分間まさに機関銃のようにしゃべりまくる発表となった。

1999年に林崎がコールドスプリングハーバー研究所で完全長cDNAのデータベースとクローンバンクについて発表すると、世界中の大学や企業から、分与・供与の要請が理研に殺到した。特に、ES細胞の完全長cDNAデータベースがあることが知れ渡ったので、iPS細胞の開発目的で、現在の京都大学教授山中伸弥

のiPS細胞開発と競合していたグループからの共同研究の申し入れや、使用権を購入する申し込みなどが頻繁に送られてきた。研究室の中でも、このような類似のプロジェクトを起こす意見も出た。他機関に分与できる科学的産業的価値を持ったプラットフォームの取り扱いについて、和田昭允GSC所長や理事会にも相談したが、理研内に経験者は皆無で、その取り扱い基準や基本ポリシーをつくる作業は、結局自分たちでやらねばならなかった。

さらに、次に述べる日本の特許庁の歴史的活動が、理研のポリシー決定に大きな影響を与えた。そしてそれが、その後の理研のオミックス科学の基本方針を決定的なものとし、FANTOM (Functional Annotation of Mammalian cDNA and Genome) の国際コンソーシアムの結成という方向に進んでいったのである。

第2節 国際FANTOMコンソーシアムの結成

日米欧の特許庁3極会議と公開ポリシー

1995 (平成7) 年に、尾身幸次衆議院議員 (のち財務大臣) らの努力によって科学技術基本法が制定され、日本の科学技術は大きく飛躍することになった。この科学技術基本法制定の引き金となった背景の一つに、ゲノム特許の問題があった。アメリカのベンチャー企業がcDNAの部分配列 (完全長cDNAではない) であるEST (発現配列タグ) を大量にシーケンスし、その配列全ての特許を取り始めたのである。EST配列の特許が認められれば、未来に開発される全ての薬剤 (タンパク製剤、RNA製剤はもちろん、低分子薬剤もその標的となるタンパク質まで) は、全てこのEST特許の中に入ってしまうため、未来の創薬産業界に大きな影響を及ぼすことになる。アメリカ、欧州、日本の3極の特許庁の間で、ヒトのゲノムの配列やRNAの配列が特許として認められるべきかどうか、議論が沸き上がった。

理研ゲノムプロジェクトでも、自分たちが出した完全長cDNAのシーケンスの知財をどのようにすべきか、重大な決定をする時期に差しかかった。結局、「防衛的にでも特許を出して、後で取り下げればよい」という考えで、とにかく出た配列 (タグ配列) だけでも、とりあえず、出願することになった。1999年のことである。当時の特許庁は、出願はまだ紙ベースで、電子的に出願できなかった。そのため、当プロジェクトの特許担当の渡辺幸彦らが、シーケンスが記載してある明細書を詰めた段ボール箱を台車に満載し、特許庁の窓口を持っていった。前代未聞の出願スタイルであった。その後すぐ、2000年1月より、特許庁の出願システムは電子化され、オンライン手続が開始されたのだった。

日本国特許庁は、3極会議に臨むため、ゲノムDNAやRNAの配列情報に特許を与える根拠があるかどうか見解をまとめる必要があった。1998年1月より特許庁の特許技監に着任した佐々木信夫は、ゲノム分野の研究者から本件に関する意見の聴聞を行った。上述のように林崎は多くの完全長cDNA配列の特許出願をしていたこともあり、佐々木から意見を聞かれた。確かに、理研自らは多くの

配列の出願をしているものの、冷静に考えると、少なくとも自然界に元から存在するゲノムの配列そのものは産業発明にならないという基本概念もあり、林崎は特許にするべきでないという意見を述べた。

そもそも特許概念に反すること、さらに、その影響の大きさから、国際特許3極会議は、日本と欧州がアメリカを引っ張る形で、遺伝子配列そのものの情報を特許として認めないという結論に至った。前述のように、この世界の流れが、全てのRNA分子の完全長配列を突き止めようとする理研のゲノム研究の基本的方針に、決定的な影響を与えた。日本の国益を守るために遺伝子配列の特許を出す意義がなくなったため、論文発表前でもアカデミアに開示できるようになった。このことはプロジェクトの運営上、非常に大きな意義を持ち、配列データをアカデミアで共有するという基本概念が定まったのである。

完全長cDNAデータベースは、①個別の使用権を販売する、②個別の共同研究者のみに使用させる、③公開する、といった選択肢があったが、理研の使命として、きちんと機能注釈をつけた学術的に完ぺきな形のデータベースを、論文発表と同時に無料公開することとした。この決定は、理研が開発したこれらのリソースは日本の国税を使って開発したものであることを考慮したものであり、さらに、林崎の上司である和田が常々GSCのメンバーに説明していた「常に公正であることが重要」という教えに従ったものである。

個別研究にデータベースを見せたり、クローンを使ってもらうだけの個別の共同研究や、金銭ベースでの使用権分与の依頼などは全て断り、論文発表に全力を注いで一日でも早くデータを公開することを目指した。この決断によって、次に述べる国際FANTOMコンソーシアムを設立運営するために必要不可欠の環境が整い、コンソーシアムのポリシーが確立したのである。

世界から「いっしょにやろう！」と声がかかる

1999年のコールドスプリングハーバー研での発表以降、完全長cDNAのデータベースとクローンバンクが理研だけにあることが世界中に知れ渡った。がん遺伝子の発見でノーベル医学賞を受賞した当時のアメリカのNIH長官のバーマス(Harold Varmus)がこれを聞きつけ、NIHは第三者機関として、完全長cDNAクローンの機関内の試験的配布とデータベースの検証を行うという共同研究契約を理研と結んだ。

この契約に調印した理研ゲノム科学総合研究センター(GSC)所長の和田は、「NIHみたいな一流の研究機関と共同研究することで、『林崎が作ったシーケンスは乱数表で出している』とか、『林崎がつくったシーケンサーは空の箱だ』といった根も葉もない誹謗中傷を黙らせることができる」と予測した。ところが、2000年5月8日の『朝日新聞』の朝刊第一面で「理研、事実上無償供与」という見出しの記事を書かれてしまった。つまり、「日本国の宝を勝手にアメリカに無償供与した」というニュアンスを匂わせたタイトルには、さすがの和田も呆れ果てたのだ。ともあれ、理研の完全長cDNAは、「乱数表による偽造データ」から「日本国の宝」へと出世したのである。

2000年4月14日号の*Science*誌に、マウス遺伝学の科学史年表を含む特集記事が掲載された。この中で、1664年にフック（Robert Hooke）がマウスを実験科学目的に使用したのを手始めに、節目となる発見や研究成果が45項目記載された。その最後の項目に「1999年、日本の林崎のグループは、最初の完全長cDNAの配列を決定した。2万個のDNAが遺伝子発現の解析のためにマイクロアレイにプリントされた。NIHは所内の科学者がその全データベースにアクセスできるようにしており、他の研究者も同様の機会が得られるよう希望している」と書かれた。林崎らの完全長cDNA研究の活動が世界に認められた証であった。年表の最終行は「2000年、マウス遺伝学が離陸した」である。

スタンフォード大学のブラウン（Patrick Brown）からも、共同研究の申し入れがあった。完全長cDNAは、ブラウンが開発したcDNAマイクロアレイの絶好の材料になる。これらのcDNAを機関内試験配布して検証することを目的として共同研究を結び、理研の完全長cDNAが確かに機能することが検証された。NIHやスタンフォード大学との共同研究の成功は、理研の完全長cDNAとデータベースが科学的に正しいことの証明書となり、次のステップとして、機能注釈をきっちりつけ、論文公開に運ぶ段階に進むことができたのである。

研究者コミュニティからも非常に多くのアプローチがあった。ある日、オーストラリアのクイーンズランド大学のヒューム（David Hume）から突然、理研ライフサイエンス筑波研究センターの林崎のオフィスに電話が入った。理研の持っているデータに機能注釈付けをしていくプロセスを一緒にやらないかという提案であった。林崎らも同じことを計画していたために、共同研究者として加わってもらうことになった。ヒュームは「本件は重要だから、今からそちらに行く」と言って電話を切った。翌日林崎が出勤するとヒュームが部屋の前に座っており、「今からそちらに行く」の意味がその時に分かったのだった。特許問題がなくなり、ヒュームの参加や和田の強いサポートもあって、次に述べる国際FANTOMコンソーシアムの結成に至ったのである。

第1回FANTOM会議は2週間の泊まり込み合宿

当時理研が持っていたデータベースは、塩基配列が記載されているだけで、これだけでは、実際のルーチン使用には程遠いものであった。一つ一つの塩基配列に注釈付けを搭載した世界全体が使える国際標準データベースを作成し、さらにそれらを記述した論文をつくり、国際標準仕様のデータベースを公開維持するシステムを構築しなければならない。そのためには研究者コミュニティの協力が必要不可欠であるため、林崎らは、国際FANTOMコンソーシアムの設立を決めた。多くの支援が得られそうなので、早速、データをコンソーシアムで共有する



第1回FANTOM会議
(2000年8月、理研ライフサイエンス筑波センター)

ための準備を始め、第1回国際FANTOM会議を2000年8月、理研ライフサイエンス筑波研究センターで開催する運びとなった。

会議は合宿形態で、一流の科学者が2週間泊まりこんで、アノテーションを行うことになった。もちろん、それまでRNAの配列の機能注釈を付けた例などなかったため、全て、その注釈付けのシステムそのものを自分たちで作らねばならなかった。最初の1週間は、何をしてもよいのか全く分からず、毎日朝から晩までどのようにするかの話し合いを重ねた。既知のタンパク質配列との類似度を見ることで、その遺伝子はどのような機能を持っているのかを予測したデータベースになる。それをどのように行うのか、一日議論してはそのプログラムを一晩で作し、翌朝から、解析したデータをさらにどのように処理するのかを決め、また一晩のうちにプログラムを作り直し、解析データを翌朝までに作るという、徹夜の連続で注釈付けしたデータベースを作り上げた。

このほぼ1週間、不眠不休のプログラム作成とコンピュータ解析を、坊農秀雅（現ライフサイエンス統合データベースセンター特任准教授）、岡崎康司チームリーダー（現順天堂大学大学院医学研究科教授）たちがやり遂げた。会期2週間の内、最後の数日で一気にデータベースの注釈付けを終えることができたのである（図1）。

ヒト全ゲノムのドラフトより1週早く*Nature*に発表

このようにして、世界で初めての高等動物のRNAの百科事典ともいえる国際標準の完全長cDNAクローンとデータベースが完成した。理研の完全長cDNAデータは、当時並行して進んでいたHUGOによるヒト全ゲノムドラフトシーケンスの論文の中で、ゲノム配列からRNAになる部分（エクソン）をコンピュータ予測した仮想的遺伝子の実在検証に使われた。そのため、林崎らのデータが先に出版されることが必須であり、*Nature*誌とHUGOとが話し合った結果、2001年2月15日号のヒトゲノム論文の一つ前の2001年2月7日号の*Nature*誌に、理研と理研がリードする国際FANTOMコンソーシアムが開発した2万種類の完全長cDNAクローンと国際標準完全長cDNAデータベースが発表されることになった。

論文の発表と同時にデータベースが公開されることが事前に世界中に知れ渡っていた。そこで一刻も早いデータへのアクセスを希望している組織があることは分かっていた。*Nature*の発表はアメリカ東部沿岸時間で2001年2月7日午後2時、つまり日本時間の2001年2月8日午前4時であり、理研は同時刻にデータベースのパスワードを解除した。すると、あるアメリカ企業は、その1秒後からダウンロードを始めたのである。これを見て、ビジネス魂の迫力にスタッフは呆気にとられたのであった。

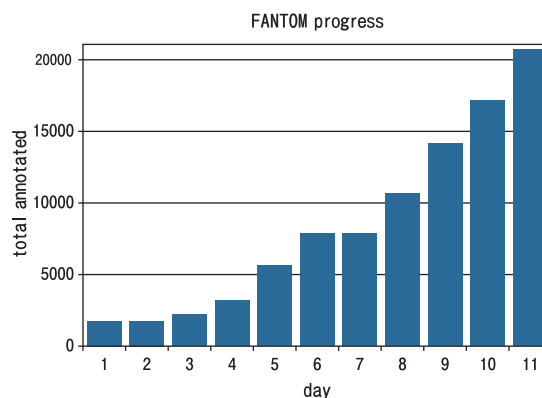


図1 会議期間中に注釈付けされたマウスcDNAの推移



インパクトの高い成果を出したFANTOM2
(2001年4月、理研横浜研究所)



図2 新たな技術を駆使してRNA大陸を発見

RNA大陸の発見—FANTOM2からFANTOM3へ

マウスのさまざまな組織から取った1800万個の完全長cDNAを分類し、10万5000種類の完全長cDNAについて全長シーケンスを解明した研究をFANTOM2（6万個）とFANTOM3（10万5000個）で行った。これらのシーケンスデータに機能注釈を付けた国際標準データベースを作るため、横浜で開催した2001年4月FANTOM2会議を皮切りに、これまで合わせて10回の国際FANTOM会議を開いた。

このFANTOM2の完全長cDNAは、国際マウスゲノムコンソーシアムが解読したマウス全ゲノム配列とともに、2002年12月7日号の*Nature*特集号で発表された。これにより、マウスが、ゲノムと完全長cDNAの両方が解析された（RNAのゲノム上の位置が解明された）最初の生物となったのである。

FANTOM2のデータはいろいろな意味で非常にインパクトがあった。そもそも遺伝子とは、エクソンを共有するRNAを合成するゲノムDNAの領域と定義されていた。しかし、FANTOM2のデータにより、多様なスプライスバリエントが発見され、従来の二つ以上の遺伝子のエクソンを共有する新しいスプライシングバリエントが見つかる、それまで二つ以上の遺伝子だったものを一つの遺伝子としてカウ

ントせざるを得なくなった。こうして新しいRNAを発見すればするほど、遺伝子の数が減るという予想外の現象に遭遇し、従来の遺伝学で定義されてきた遺伝子の概念そのものが機能しなくなった。現に、それまであった「ヒトゲノムに遺伝子はいくつあるのか」という議論は、いつしか消えていったのである。

さらに驚くべき事実が明らかになった。予想以上の数のncRNAが発見されたのである。FANTOM2は、単に完全長cDNAの新しい配列を求めたという意義を持つだけではなく、種類別のRNAの過半数が、タンパク質をコードしない非コードRNA（ncRNA）であることを明らかにしたのである。この発見は、それまでの現代生物学の基本とされてきたセントラルドグマ（中心教義）を覆すものであった。すなわち「遺伝子DNAから作られるRNAはタンパク質をコードするものである」とは、もはや言えなくなったのである。これが「RNA大陸の発見」であった（図2）。

ワトソン・クリックのDNAの2重らせん構造の発見から50年間、世界のライ

フサイエンス研究者は、ほとんど全て、タンパク質をコードする遺伝子だけに焦点を当ててきた。ところが、RNAつまり遺伝子発現の機能面から見てみると、遺伝子の大部分がこれまで研究対象から外されてきたという、まさに驚くべき事実を突き付けられたのであった。FANTOM2のこの結果は、学会関係者の間でも、すぐには受け入れられなかった。しかも、細胞内の存在量が少ないRNAにncRNAが多く存在していたため、FANTOM2の発表戦略を話し合う過程で、内部のスタッフからも異論が出たのである。「世の中の常識としては、RNAとはタンパク質をコードするものである」とされているし、「ゲノムDNAから漏れ出てきた転写物ではないのか」という疑いも出され、データの中に入れて発表するべきではないかという声上がるほどであった。プロジェクトの統括である林崎は決断を迫られた。



「RNA新大陸の発見」について記者会見（2005年9月2日）

賛否両論の中、林崎はありのままのデータを発表することを即決し、実際、それがすぐに実行された。案の定、FANTOMのncRNAは、タンパク質コードRNAと比べて、マウスとヒトの種間の保存がされていないことを理由に、単にゲノムDNAから漏れ出て転写（合成）されたものではないか、と推測する論文まで出版された（Jun Wong et al., *Nature*, Vol.431, October 14, 2004）。さらに、*Nature*誌上や学会での議論を見て、後発の完全長cDNA プロジェクトの中には、5'端からの塩基配列でタンパク質のコドンが出てこないncRNAと考えられるクローンを捨てて解析する、という戦略を取るプロジェクトさえ現れ、それが良識あるプロジェクトであるかのように扱われた。

この議論は、FANTOM3の発表によって急展開した。それまでタンパク質をコードしている領域、つまり機能している領域は全ゲノムのわずか2%であったのに、ヒトゲノムの70%の領域でncRNAが転写（合成）されていること、生体内での72%のRNAには、逆鎖のDNAから読まれるアンチセンスRNA（ほとんどがncRNA）のパートナーが存在すること、それらが転写レベルの調節に機能し合っていることが、FANTOM3の論文の骨子であった。

林崎は、*Science*誌のゲノム担当編集長であるジャスニー（Barbara Jasny）に会いにワシントンDCに出向き、FANTOM3の主論文出版について相談・助言を受けた。そして、他にもFANTOM2で発表したncRNAのさまざまな生物活性を証明した論文が投稿されているので、これらとともに*Science*誌がRNA特集号を組んでくれることになった。

理研が単離し、国際FANTOMコンソーシアムと理研が機能注釈を付けたncRNAの生物学的活性を証明した他グループからの6報の論文とともに、理研がリードしたFANTOM3の主論文2報、計8報が、2005年9月2日号の*Science*誌に発表され、同日、市ヶ谷にあるアルカディア市ヶ谷で記者会見を行った。生

物学の常識を変える発表の後、1995年よりポスドクとして理研グループに参加し、今回の主論文の一つの第一著者となったイタリア生まれのカルニンチに声をかけた。林崎が「You should be proud.」という、うれしそうな顔が返ってきたのである。

振り返ってみると、理研はFANTOM2の完全長cDNAクローンの頒布を、発表後すぐに開始し、データベースを無料で公開していた。このことが、ncRNAの生物学的意味についての議論の行方を決定付けた。理研自らもあまたあるncRNAの機能の一部を解明したが、全世界のさまざまな研究室が、理研が開発したFANTOMのリソースを使い、さまざまなncRNAが生物活性を持つことを次々と解明してくれていたのである。自分たちのリソースを他のグループに無料で開放することが、結局は、自分たちを救ってくれる結果となった。

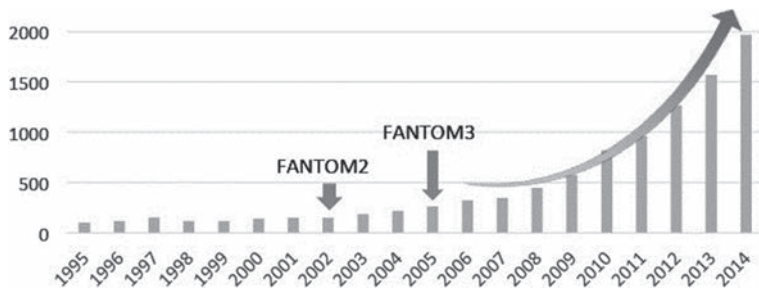


図3 FANTOM3以降、ncRNAの新機能に関する論文は指数関数的に増加した

その後、FANTOM2と3の発表の後、それまで理研のFANTOMデータを真向から否定し続けてきたグループは、沈黙するどころか、自らncRNAの新しい機能を発見する世界競争に加わり始めた。それもあって、図3に示すように、世界のncRNAの新機能を示した論文は、指数関数的に増えていった。このよ

うにFANTOM2と3の活動は、全く新しい研究の世界を切り開いていったと言える。

世界のスタンダード技術に

マウスやヒトの完全長cDNAの国際標準データベースは、その後、ほかの生物の完全長cDNAデータベースに拡大拡張した。そうしたライフサイエンスの基盤を整備する活動は、研究コミュニティからの要求でもあった。例えば理研植物科学研究センターの篠崎一雄センター長のリクエストによるシロイヌナズナ完全長cDNA、農水省の委託によるイネ完全長cDNA、イリノイ大学のロビンソン (G. E. Robinson) との共同研究によるミツバチ完全長cDNA、ルービン (Gerald Rubin) のリクエストによるショウジョウバエ完全長cDNA、ハーバード大学のカーシュナー (Marc Kirshner) のリクエストによるアフリカツメガエル完全長cDNA、そのほか、ポプラ完全長cDNA、大豆完全長cDNAなど、種々の生物の完全長cDNAバンクが次々と共同研究により解析され、現在では、各分野でライフサイエンスの基礎インフラになっている。さらに、これらの完全長cDNAクローンバンクも研究コミュニティの要望が多く、頒布活動が必要となってきた。そのため、当時の理研ベンチャー制度によりクローン頒布を目的として (株) ダナフォームが設立され、クローン頒布を開始した。

もともと、 -80°C で保存し、ドライアイス詰めで送付することが基本となっているクローンバンクは、海外などにも送付する必要があり、コストもかかる。そ

のために、林崎らは“DNAブック”という新しい技術を開発した。水溶紙に精製した完全長cDNAの実物をインクとして点着印刷し、室温普通郵便で送付、その後、エンドユーザーは、該当する水溶紙の部分をパンチアウト、完全長cDNAを含んだ水溶紙ディスクをPCR反応液に入れて増幅する。紙は反応液に溶けて、DNAは増幅して出現するという仕掛けである。これにより、マウスの10万種類の遺伝子（完全長cDNA）が印刷されている本、『マウスエンサイクロペディアDNAブック』が出来上がった。『ヒトcDNAのDNAブック』や、上述した『イネcDNAのDNAブック』、さらに、神奈川県水産技術センターとの共同研究によるヒラメなどの『アクアDNAブック』が次々出来上がった。こうした印刷物は、単に情報を運ぶだけでなく、物質を運ぶ媒体としても使われており、その最初の例として意義があるため、横浜市立図書館、神奈川県（松沢成文知事）、日本科学未来館（毛利衛館長）などに寄贈された。



『DNAブック』を日本科学未来館（毛利衛館長）に寄贈（2003年11月5日）

完全長cDNAクローンはタンパク質を発現する道具としても有用である。そのために、タンパク質コード領域だけをさまざまな発現ベクターに組み込めるシステムを作るプロジェクト（ORFeomeプロジェクト）が、NIHを中心とする国際ORFeomeコンソーシアムで推進されることとなった。メンバーは、ハーバード大学のヴィダル（Marc Vidal）グループがアメリカのグループをまとめ、理研の林崎グループも完全長cDNAとORF発現クローン構築を行い、日本より参加した。このプロジェクトにより国際標準ORFeome資源バンクが作られ、全世界に頒布されている。

転写ネットワーク解析への追い風

FANTOMプロジェクトは「ゲノムに何が書かれているのか？」という大きなテーマを目指して進められてきた。完全長cDNAの配列は、明らかにその第一歩であった。受精卵から個体に至り、さらに生殖細胞系列から次の世代をつくる、というライフサイクルを分子レベルで理解するには、ゲノムに書かれている遺伝子がどのような順番で、どのようなメカニズムで制御され、ライフサイクルに沿って発生、分化、老化などが進むごとに変化していく遺伝子調節のダイナミズムを解析する必要がある。

遺伝子DNAがRNAに書き写されるためには、RNAが5'端から書き写し始められる部分（転写開始点）のすぐ上流に、直接的に転写の頻度を決定する（RNAポリメラーゼを呼び込む）プロモーターとよぶ領域が存在する。全ゲノム配列は明らかになったものの、どの部分がこのプロモーターであるのかは、依然として全く未知であった。

CAGE (CAP Analysis of GENE Expression) RIKEN original technology to reveal "GENE REGULATION"

CAGE analyzes 5'-end of the capped transcripts by DNA sequencing.

- 1) Precise transcriptional starting site (TSSs) are clarified.
- 2) Expression profile at each promoter (not gene9 could be analyzed).

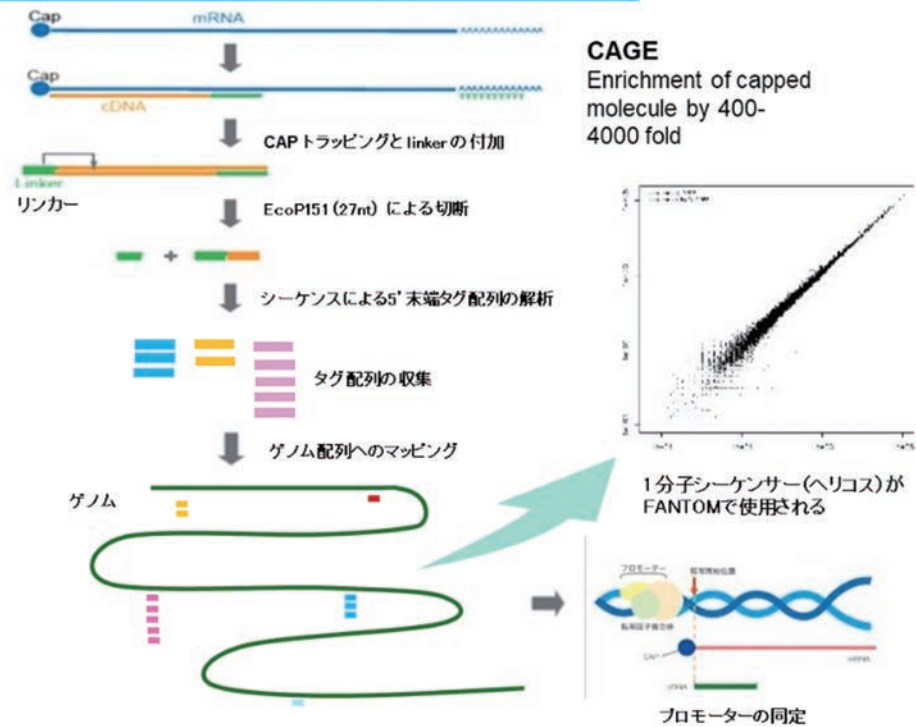


図4 理研で開発されたCAGE法。FANTOMでもこの技術を使って多くのデータを輩出することができた

林崎らは、この転写開始点を明らかにするために、CAGE法 (Cap Analysis of Gene Expression) という技術を開発した。CAGE法は、RNAの5'端のcDNAのみを集めて、シーケンサーで5'端のタグ配列のシーケンスを行い、ゲノム上にこれらのタグ配列をマップすることによって、ゲノム上のどの位置から、どちら方向に、どの細胞で、いつ、RNAが発現しているかを、1塩基の解像度で解析することができる。このCAGE法は理研オリジナルの手法であり、現在でも遺伝子のプロモーター活性を見るのに、世界で現存する唯一の方法である (図4)。

そして、このCAGE法にはさらなる追い風が待っていた。次世代シーケンサーの登場である。次世代シーケンサーは、RNAの5'端タグ配列を、大量に、しかも、安価に決定することができる。CAGE解析の精度は解析するタグ配列の数に依存するため、次世代シーケンサーの登場により、CAGEの精度を飛躍的に向上させることが可能になった。

その一方で林崎は、次世代シーケンサーで配列を求めるDNA検体の調整に用いられているPCR法が、タグ配列の数にバイアスを生じさせており、解析の再現性を悪くしていることに注目していた。CAGE法の要と言えるタグの数にバイアスを生じさせないためには、PCRを使わない次世代シーケンサー、1分子シーケンサーが必要である。林崎は、世界で初めて1分子シーケンサーであるHelicos

社のシーケンサー (Heliscope) が、ボストンで開発されていることを見つけた。2006年、当時Helicos社のCEOであったロンバルディ (Steven Lombardi) に連絡をとり、共同研究契約を結んだ上、ボストンに伊藤昌可を派遣した。伊藤がボストンで出したデータの再現性は、相関係数が0.999という驚異的な見事さであり、HeliscopeはCAGE法にとってまさに渡りに船の技術であった。

マウスやヒトの細胞中には、遺伝子からRNAが読まれる時に調節を行っている転写因子が約2000種類存在する (ヒトゲノム中には約2000個の転写因子遺伝子がある)。また膨大な数の非タンパク質コードRNA (ncRNA) が存在する。細胞がその表現形質を保つためには、その細胞の種類に特異的な複数の転写因子どうしが正負双方向の制御を行っており、ネットワークを形成しているはずである。このため、細胞がその表現形質を保っている以上、ネットワークを形成する転写因子の細胞内濃度は、一定のはずと考えてよい。この濃度が一定の基準より低くなったり高くなったりして発散すると細胞はそれまでの転写ネットワークを保つことができなくなり、次の (別の) 種類の細胞へと変化する。

これらの転写ネットワークを実際に解析解読するプランは、林崎の中で2003年ごろから2005年にかけて構想がまとまった。特に、前述した大量のncRNAの発見により、これらのncRNAもまた転写ネットワークに関与していると考えられるので、解析対象がさらに拡大した。2004年9月、FANTOM3の会議の席で、林崎は転写ネットワークを実際に解析解読したいという夢を全員の前で語ったが、メンバーのほとんど全員は、「そんなものできるはずがない」と笑い出してしまった。しかし、それ以後のFANTOMの方向性は、林崎の中では明瞭であった。

第3節 夢の実現—国際標準から医療応用へ

FANTOM4-FANTOM5

林崎が構想していた転写ネットワーク地図の概念図は、CAGE法により、実測データに基づく転写ネットワークとして解明することが可能になった。CAGE法で遺伝子の転写を調節しているプロモーター活性を計測すると、転写因子をコードする転写因子遺伝子間の調節のネットワーク地図を描くことができる。そのためには、非常に高い再現性と精度を持つシーケンサーが必要であったが、前述のHeliscopeで出したCAGEデータ (Heliscope CAGE) は、転写ネットワークを解析するのに理想的なシステムとなった。

実際に、ヒト単芽球白血病細胞株がPMA (ホルボールミリスチン酸) で単球に分化する系を題材に、CAGE法を用いて、実データに基づく細胞の転写ネットワークを解析するシステムを作ったのが、FANTOM4である。このFANTOM4のトライアルは、転写ネットワークを描いた世界最初の例であった。ヒトの細胞の中にあるプロモーターは有限個であり、例えば、単芽球が単球に分化するのに、2万9857個のプロモーターが活動していることが判明した。こうして、ヒトの細胞の分化を分子レベルで理解できるまでの距離が見えてきたのである。

転写ネットワークは、細胞の表現形質を決定論的に決めるものである。ヒトの体を構成するあらゆる種類の細胞の転写ネットワークを調べることは、さらに、その細胞で活動しているプロモーターを明らかにすることは、細胞の形質を理解する上で非常に大切である。組織学の教科書には、人体を構成する細胞種の数として、ほぼ400種が紹介されており、これらの細胞種の転写ネットワークを全て明らかにし、ネットワークの百科事典を作ることは、その後のライフサイエンスの基本的なデータとなり得る。そこで、林崎グループのフォレスト (Alistair Forrest) は、人体から単離される単一種類の細胞を集め始めた。FANTOM5では、ヒトをはじめとする数種の哺乳類の細胞や細胞株、組織1000種 (そのうち、初代培養細胞180種) について、全てCAGE法でプロモーターの位置やその活性を明らかにした (これらは全て共同研究者から提供されたものである)。その結果、ヒトゲノムの中にプロモーター活性場所が18万カ所見つかり、国際標準のプロモーターアトラスとなった。

さらに、このデータベースから、次への跳躍となる発見がなされた。2006 (平成18) 年12月まで林崎のセンターのユニットリーダーであったサンデルイン (Albin Sandelin) (この後、コペンハーゲン大学の教授に就任) が、理研の国際共同研究FANTOMの活動の一環として、遺伝子プロモーターの活性 (遺伝子の転写活性) を遠隔から制御しているエンハンサー領域を、系統的なシステムを用いてヒトゲノム中に4万4000個発見した。エンハンサー領域には双方向性にプロモーター活性 (転写開始点) があることを利用し、上述の大規模かつさまざまな種類の細胞種について解析されたプロモーターアトラスの中で、双方向性プロモーター活性を持つ領域を抽出することにより、成し遂げることができた。

エンハンサーから転写されるこの双方向性のRNAは、エンハンサーRNA (eRNA) と名付けられた。eRNA (エンハンサーの活性) の発現は、組織・細胞の種ごとに特異性が非常に高いこと、また、種を超えて保存されている配列を持つエンハンサーは、組織特異性や種を超えて保存されていることを遺伝子導入動物で証明することができた。さらに、このエンハンサー領域にある変異は、特定の疾患と相関を示すことが非常に多いことが判明した。ヒト疾患の遺伝子変異の80%という大多数が、遺伝子領域外 (もちろん、非タンパクコード領域) であるエンハンサーとプロモーター領域に存在することが判明したわけで、これは非常に驚きであった。これらのエンハンサーとプロモーターの論文は、2014年に*Nature*に2連報 (Vol.507, 455-461ページとVol.507, 462-470ページ) として、2015年に*Science* (Vol.347, 1010-1014ページ)、2017年に*Nature* (Vol.543, 199-204ページ) に報告された。当然、ライフサイエンス全ての分野で用いられるデータベースとなり、2014年、2015年の2年間で、FANTOM5の2報の論文の被引用インデックスは日本で1位と2位を占めた。

特に、プロモーターの論文は、261人も共著者を有する論文となり、双方の論文で作られたプロモーターアトラスとエンハンサーアトラスは、国際標準のデータベースとなるとともに、FANTOM活動の医療応用への道を開くものとなった。すでに述べたように、これらの転写ネットワークは、細胞の表現形質を

決定論的に決めている。したがって、今回解析したネットワークを別の細胞に移植すれば、iPS細胞を経ることなく、直接目的の細胞を作ることができる。実際、そのために、どんな細胞にどの転写因子遺伝子を導入すべきなのか、FANTOM5で作ったネットワークデータベースの地図から推定することができる。2016年、そのためのコンピュータソフトMogrifyを発表し (*Nature Genetics*, Vol.48, 331-335ページ)、さらに、Webサイトに公開した。

FANTOMコンソーシアムの軌跡

以上、FANTOM4とFANTOM5は、新しいRNA（完全長cDNA）の発見や配列決定のみならず、プロモーターや従来知られていなかったエンハンサーなどの遺伝子の調節領域を明らかにし、トランスクリプトームの研究領域に飛躍的発展をもたらした。

FANTOMデータは、2016年5月現在で、3分間ごとに1回アクセスを受ける国際標準のデータベースとなった。世界のコミュニティーが使い続けているのである。FANTOMの活動は、国際的に高い評価を受け、数々の著名誌に紹介記事が掲載された。例えば、2000年の*Nature*誌では、ヒトのモデル動物であるマウスを材料に、あらゆる組織、発生段階における完全長cDNAの塩基配列を決定し、そのデータおよび完全長cDNAクローンを世界に向けて公開したFANTOMの活動が高く評価された。当時進んでいたゲノムプロジェクトと合わせ、ライフサイエンスの重要な基盤と位置付けられた (*Nature*, Vol.407, 279ページ)。

また、2002年の*Nature*誌では、FANTOMのマウス完全長cDNA 6万6000種の配列解析は、研究者の遺伝子塩基配列決定や構造の解析に費やされる時間と手間を省き、機能解析を促進する強力なプラットフォームになると評価された (*Nature*, Vol.420, 456-457ページ)。さらに、2010年に*Science*誌が選んだ過去10年間の10大研究の一つとして、機能性非タンパクコードRNAの発見がその後ノーベル賞に輝く山中のiPS細胞とともに選ばれた。その中で、FANTOM3の成果 (P. Carninci et al., *Science*, Vol.390, 1559-1563ページなど) は、非コードRNAが遺伝子発現制御において重要な役割を果たすことを見出した研究、として紹介された (*Science*, 330, 1614 (2010))。

FANTOMのように、日本が過去22年間常に世界をリードしてきたコンソーシアムは、ほかに例がない。FANTOM5だけでも、500人以上、20カ国からの参加者（2015年現在）があり、FANTOM1から通算で1100人以上、114機関、25カ国からの参加者となる。FANTOMコンソーシアムは、世界最長の歴史をもつ世界最大の参加者を誇る国際コンソーシアムとなった。これらのFANTOMの活動により林崎をはじめ、数多くのFANTOMのメンバーが受賞している。しかし、2016年オーストラリアEURECA賞は、個人にではなく、FANTOMという国際コンソーシアムに対し、授与されたことは喜ばしい限りである。ライフサイエンスのコミュニティーに対し、多大な功績を認められた国際FANTOMコンソーシアムを22年の間リードしてきたことは、理研が、そして日本が、世界へ貢献してきたことのVisibilityを示すものである (図5)。

FANTOM History

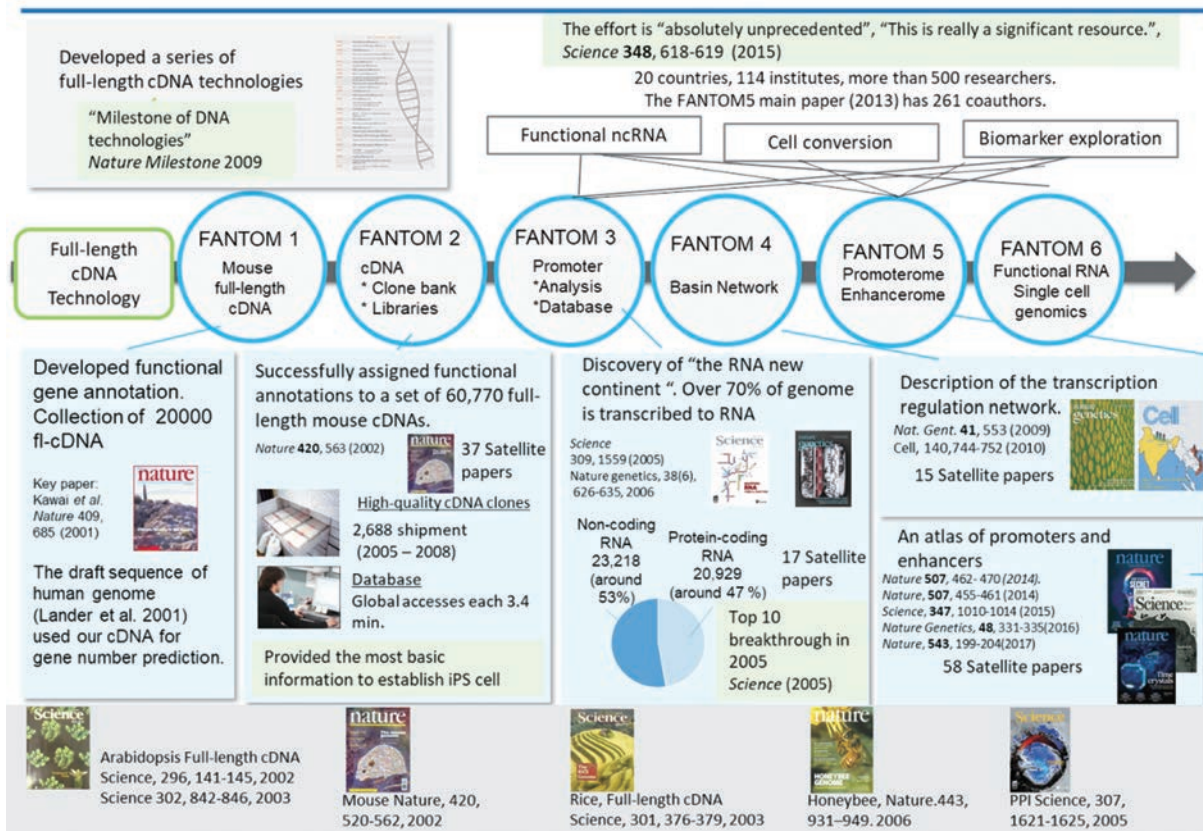


図5 FANTOMの歴史 Full-length cDNA technologyからFANTOM6まで、日本がコンソーシアムをリードし、多くの成果を上げている。

FANTOM1のスタート当時は、このプロジェクトの目標は、基礎科学に貢献するデータベースとクローンバンクの作成であった。しかし、研究成果のバトンリレーにより、FANTOMデータベースを利用して山中がiPS細胞を発明し、さらに、理研CDBの高橋政代によるiPS細胞誘導網膜色素上皮が作られ、移植手術に用いられるまでになった。基礎科学を支える国際標準データベースが社会実装にまで至る好例といえる。これらは全て日本オリジナルの誇り高い成果である。それはとりもなおさず理研が100年間常に追いつづけてきた道にほかならない。