

情報の普遍的概念を求めて (1) : 遺伝情報系と言語情報系

Searching for a Universal Concept of Information: Part 1. Genetic and Linguistic Information System

竹村 彰 祐*

Shosuke Takemura

Abstract

As an approach from natural science in searching for a universal concept of information, genetic information system was compared with linguistic information system. The word 'genetic information' appeared before 1960, and its concept already in the late 19th century.

DNA, the major component of chromosome, consists of four kinds of bases and carries genetic information on protein structure and function. Each protein molecule is formed by chemical linkages of many amino acids. A set of three contiguous DNA bases named codon codes for an amino acid. The codon sequence of a gene specifies the amino acid sequence of a protein molecule.

Several structural steps, highly analogous to the language system, were seen in the genetic information system; DNA bases, codons, genes, and chromosomes, corresponding to phoneme(letters), words, sentences, and books. Expression of genetic information was schematically closely homologous to that of usual language. Results of other comparisons also emphasized that the genetic information system can be regarded as a sort of the natural linguistic information system. The genetic language system is inferred to have evolved much more than the usual language system.

竹村彰祐* : 愛知淑徳大学教授 (文学部図書館情報学科)

JOURNAL OF LIBRARY AND INFORMATION SCIENCE, Vol.10, p. 1-14(March 1997)

本論文の内容は、愛知淑徳大学図書館情報学会 (1995年11月16日) における講演を基にしたものである。

はじめに

本論文は、自然科学的な考察によって情報の普遍的概念を探るのを目的とする。自然科学は経験や観察を重んじ、それによって得られる事実を重視し、個々の事象から全体に通用する一般則を導き出す。そうした手法によって発展してきた自然科学は、普遍的な情報概念を導くのに適しているであろう。著者は遺伝情報の伝達・発現という分野で長年研究してきた。したがって情報を普遍的な視野で捉えるのに、分子遺伝学を足場にして、生化学ないしは分子生物学、さらには広く自然科学から眺めようとするのは当然であろう。

古代の人は、自然は見たり聞いたり感じたりできるものとして捉えていた。たとえば、Aristoteles は自然を水と土と風と火の4つの元素(物質)でできているとした。風は空気の流れで、エネルギーが関わっているし、火は物質が酸化燃焼して、物質のもつ化学エネルギーが熱エネルギーに変わっている。しかし昔はエネルギーの概念がなかった。中世になって次第に機械文明が発達しはじめ、風車や水車の時代から蒸気機関車が発明されるまでの長い年月の間に、人々はエネルギーの概念を獲得した。そして、自然は物質とエネルギーから構成され、自然における諸現象の多くは物質とエネルギーを基盤にした関係、すなわち物理の法則、によって説明されてきた。ところが、生命はそれだけでは理解できなかった。かくして生物学は自然科学の発展から取り残されていったのである。生物は種によってそれぞれ特徴的な形態・構造を維持し、精緻に特異的に機能している。一方、地球上にある低分子の有機物とエネルギーを用いて、高度に複雑な有機構造体を合成することはできるが、それらは構造が無秩序で、ランダムで、再現性がない。物質とエネルギーだけではこうした生命の特異性は説明がつかない。それぞれの生命にその形と機能を与えるもの、つまり情報がさらに必要なのである。information

の語源はラテン語の *informare* で、*to give a shape to* (～に形を与えること) という意味である。

こうしてみると生物は実に多様な情報システムから成り立っていることがわかる。遺伝情報系、脳神経系、内分泌(ホルモン情報)系、免疫応答系をはじめ、一般的に言えば、生体外部から内部(細胞)へ、あるいは細胞から細胞へ、情報の受け渡しが絶えず行われて、生命を維持しているのである。まさに生物は情報システムの集合でできているといえる。ここで情報の授受をしているのは物質(分子)であって、言語情報や画像情報などとは違うので、こうしたシステムで「情報」の語を用いるのに抵抗がある向きもあろう。しかし、生命の情報も一般に言う情報も同じ性格を持っている。

生命に最も基本的で、これあるからこそ生命といえる性格は、自己と同じものをつくって増える「自己増殖」という遺伝的性格である。それには遺伝する情報が必要である。自己増殖は「自己組織化」能によって達成される [01]。N. J. Belkin (1976) [02] は情報スペクトルの一端に *infra-cognitive* として *heredity* (遺伝)、*uncertainty* および *perception* を挙げ、遺伝を情報の一部に位置付けている。また、B. C. Vickery ら (1987) [03] はもっと広い視点で生物における情報について触れている。本論文では、遺伝情報系から普遍的な情報概念を探るためにまず最初に自然言語情報系との対比を試み、それらの間の高度の類似性を挙げて、定義によっては遺伝情報系を言語情報系の一つと見なし得ると指摘した。

自然科学系の話について、人文・社会科学系の研究者の理解を得るのは一般に容易ではない。それで、分子遺伝学の基礎的知識のうち、論旨の展開に必要な部分をなるべく平易に解説しながら述べることにする(より詳細については文献 [04] を参照)。

1. 「情報」という語に関して

1. 1 個人の体験と情報のイメージ

人は物事を考えたり判断したりするときには、過去に経験して脳に蓄積した記憶を想起して照合するという作業をする。意識や観念や概念の形成も過去に受けた印象や経験の強さに左右される。価値観もそうであって、価値判断の基準も過去の経験の蓄積に基づいて形成される [05]。したがって、ある語彙に対する意味や解釈も個人の過去の経験から逃れることはできない。情報という言葉について抱いているイメージは、個人個人が研究してきた分野によっても違うだろうし、年齢によっても異なるだろうから、統一した情報の定義づけは難しい。情報という言葉が氾濫しているこの時代に育った若い人では、情報についてのイメージは多分著者の抱いているものよりも軽くて乾いたものかもしれないと思う。

著者の幼い頃である第二次世界大戦前から戦中にも勿論情報という言葉はあった。情報の報は知らせ（報知，報道）ということだから、「情」について言うと、情は「愛情」の情であり、「情感」の情であり、「情緒」「情念」「情操」「厚情」の情である。いわば何かひそやかな、秘めやかな、甘くて、湿っぽくて、隠れていて表に現れないようなイメージがつきまとう。ましてや「情事」や「情交」に至っては勿論のことである。情報はそうした知らせだから、例えば「耳寄りの情報がありますよ」と、そっと囁くようなイメージで受け取っていた。

中学生の頃になると、一層秘密めいた印象が濃くなったように思う。というのは、時の政府の内閣調査室には「情報部」があって、外国から日本にある出先機関に送られる外交文書や指令の分析が行われていた。指令や文書は秘密保持のために、暗号で送られてくることが多いので、その解読が主な仕事になっていた。昭和15年頃には陸軍省、海軍省の中にある調査部は、情報局に名称が変わったと記憶している。やが

て第二次世界大戦が始まると、暗号解読の成否は戦局を大きく左右した。ミッドウェーの海戦での日本海軍の敗北は、事前に米軍に暗号を傍受、解読されて、待ち伏せされていたからだということは周知の事実である。また当時の連合艦隊司令長官山本五十六の乗機が撃墜されたのも、前線視察に出かけることを事前に察知されたからであるという。日本は情報戦争に敗れたとも言えるのである。

1. 2 遺伝情報との出会い

1. 2. 1 遺伝情報という言葉の出現

終戦の年に大学に入学した著者は、化学の中でも生化学を勉強し、やがて卒業研究以来DNAやRNAという核酸を扱うようになった。核酸は生命にとって最も大事な細胞核にある物質で、染色体の主成分である。数年後に、J. D. WatsonとF. H. C. CrickがDNAの二重らせんモデルを提出し(1953) [06]、それを契機として核酸分子の構造や機能を扱う分野は急速に発展しはじめ、分子生物学と呼ばれるようになった。1950年代の末頃には、DNAに遺伝子と呼ばれる部分があって、この遺伝子が遺伝情報(genetic information)を担っている、というように遺伝「情報」という言葉を使うようになった。遺伝の情報は遺伝子DNAという物質の中に暗号となって隠されていて、解読しなければ分からない、直接見ることはできない、という点では小さい頃の「情報」について抱いていたイメージと矛盾することはなかったので、ごく自然に「情報」という言葉に馴染んでいた。

実際に文献上でいつから情報の語が使用され始めたかを特定するのはむづかしい。情報という言葉は極めて一般的であるから、主な学術論文の題目に出てはこなくても、本文に使われている可能性が高いので、全文を調査しなければ特定できない。ちなみに日本の著書に遺伝情報または情報伝達という言葉が現れるのは1962年かそれ以前である [07] [08]。Crickのcoding

の理論的研究 (1955) に端を発した遺伝暗号解読の研究の進歩につれて、また1957年頃から始まったDNA上の遺伝的メッセージのRNAへの転写のメカニズムの研究につれて、遺伝情報という語はいつしか使われるようになったと考えられる。

1. 2. 2 遺伝における情報概念の出現

遺伝情報という語そのものが文献に現れ始めた時期よりも、むしろ遺伝学の分野に情報の概念が現れた時期のほうが重要である。WatsonとCrickは、DNA構造モデル発見の当時すでに塩基配列がタンパク質の構造を決め、その結果として遺伝現象が現われるであろうと考えていた。また上述の coding problem や genetic message の語が示すように、情報の概念の出現は1960年以前にさかのぼる。

それでは、分子生物学の誕生に先行する C. E. Shannon の通信理論の研究 (1948) であるとか N. Wiener, あるいは J. L. von Neumann などによる情報科学の影響はあったであろうか。生体系、とくに脳神経系に関心を抱いていた Wiener (1948) は刺激に対する感受性は通信理論の領域に属するとして、サイバネティクスを発表したし、Neumann も自己増殖する自動機械を考えていた。情報理論あるいは遺伝暗号 (2. 3. 2 参照) の解読の研究に刺激を与えたかもしれないが、著者は当時の分子生物学ないしは生化学の論文中に情報理論が引用されているのを見たことがない。また面接と文献による非常に綿密な調査をした H. F. Judson (1979) [09] も “情報理論は、生物学の発見の経過には変化を与えなかった。……その時代の生物学的な論文 (とその中の引用文献) が情報理論のことを何もいわないことは、これを証明している” と述べている。したがって、情報理論が遺伝における情報概念の形成に関与したとは考えられない。

それどころか、1869年にDNAを発見した

Friedrich Miescher は、既に1892年の終わりに、彼の叔父であるバーゼル大学の解剖学教授の Wilhelm His にあてた手紙の中で、“数個の似ているが同じでない小さい化学物質が繰り返して連結してできた大きい分子が生物体内にあるが、「言語の語や概念が24から30個の文字で書き表されるように、この分子が遺伝の大量の通信文である可能性がある」と書いている [09]。この大きい分子とはDNAのことで、当時はこれが巨大な分子であるとは分かっていなかった。にもかかわらず Miescher のこの言及には、遺伝においては情報は伝えられるものであることが端的に示されている。情報の概念が文献の形で出現したのは、おそらくこれが最初であろう。

2. 遺伝情報の構造と機能

2. 1 DNAについて

情報というのは抽象的な存在であるが、メディアの上では記号化されたある構造として存在している。遺伝情報系の場合には、DNAやRNAが情報を乗せたり運んだりするメディアである。

細胞が増殖する前に、細胞核にある1組の染色体の各々は2倍に増え (複製)、細胞分裂によって両細胞に1組ずつ分配されるので、どの細胞も同じ数と種類の1組の染色体を含む。この染色体はDNAを主成分としているので、DNAの担っている情報はどの細胞も同一である。対になって存在している染色体の一方、つまり遺伝情報の異なる染色体の1組をゲノムという。

DNAの構造は、デオキシリボースという名の五炭糖とリン酸がつながった鎖 (紐) の糖の部分に塩基と呼ばれる化合物が結合している (図1)。塩基-糖-リン酸の結合物をヌクレオチドという。DNAはヌクレオチドが非常に長くつながった重合体である。塩基は4種類しかない (アデニン、グアニン、シトシン、チミン)。ヌクレオチドで違うところは塩基の部分だけなので、ヌクレオチドの配列を通常塩基配列といっている。

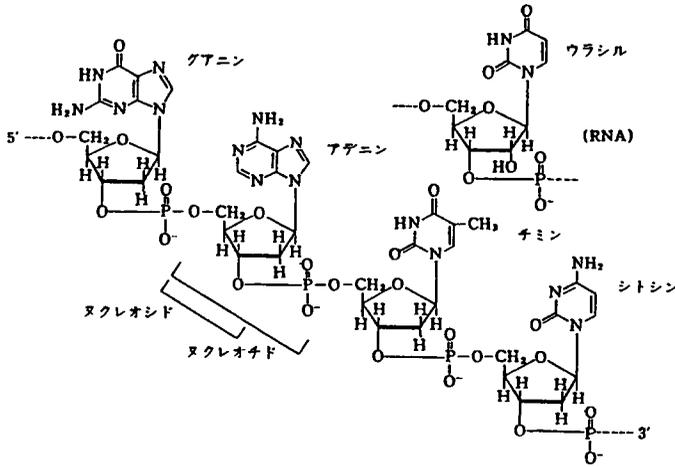


図1 DNAの構造

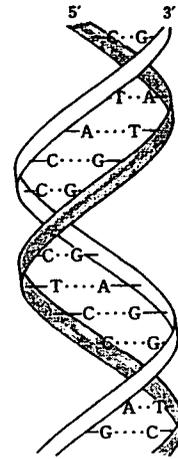


図2 DNAの二重らせん構造

生物ではDNAは2重鎖になっている。2本の鎖は互いに逆方向を向いていて、鎖の間では塩基が向き合って、アデニンとチミン、グアニンとシトシンが対を形成して、弱く結合している。この塩基対の関係はDNAの複製のときも保たれていて、どちらの鎖も鋳型になって、対になる相手の鎖がつくられるので、まったく同一の塩基配列をした2重鎖DNAができるのである(図2)。

DNAの長さは、大腸菌という下等の生物では470万塩基対、ヒトでは23対ある染色体のDNAの1組を合わせると、約30億塩基対もある。これは直線に伸ばすと約1mもの長さになる巨大な分子で、それが顕微鏡下でやっと見られる細胞核の中にコイル状になり、折り畳まれて入っている。

なお、もう一つの核酸のRNAは、DNAと比べると、糖成分がリボースで、チミン塩基がウラシルになっている。DNAの一部をコピーしてつくられるので、通常DNAよりは小さい。

によって塩基配列を表示するのは非効率であるので、欧語表記の塩基名の頭文字であるA(アデニン)、G(グアニン)、C(シトシン)、T(チミン)を用いて書き表す。すなわち、塩基配列は文字列で示すことができる。これは自然言語を文字記号で書き表すのと似ている。文章には上から下へあるいは左から右へと方向性があるように、塩基配列にも方向性がある。また文頭と文末に相当して、遺伝情報を担う部分もどこから始まりどこで終わるかが決まっている。したがって塩基物質は音声記号に対比できよう。例として、図3にヒトヘモグロビン遺伝子DNAの一部の塩基配列を文字記号で示しておく。

```

...accATGGTGACCTGACTCCTGAGGAGAAGTCTGCCGTTACTG
CCCTGTGGGGCAAGGTGAACGTGGATGAAGTTGGTGGTGAGGCCCTG
GGCAGGtgg ... 非翻訳領域 ... aggCTGCTGGTGGTCTA
CCCTTGGACCCAGAGGTTCTTTGAGTCTTTGGGGATCTGTCCACTC
CTG...
    
```

図3 ヒトヘモグロビンβ鎖遺伝子の5'末端領域の一部
小文字は非翻訳領域

2. 2 遺伝情報は文字で示すことができる

遺伝の情報は塩基の配列として存在する。DNAの構成単位である塩基という物質の構造式

2. 3 遺伝情報はタンパク質をつくる情報

アフリカに多い鎌状赤血球貧血という悪性の遺伝病は、酸素を運搬する赤血球中のヘモグロビンタンパク質のβ鎖を構成する146個のアミノ酸のうち、特定のたった1個が正常なβ鎖と異なるために起こる。これはヘモグロビン遺伝子の塩基配列のうち特定の1塩基だけが他の塩基に変化していることによる。つまり、遺伝的欠陥によるタンパク質の形成異常が遺伝的な悪性貧血として表現されるわけである。この例でわかるように、遺伝情報とはある特定のタンパク質をつくるための情報である。そして、タンパク質をつくるための情報を担っているDNA上の領域を遺伝子と呼ぶ。遺伝子は、DNA上で互いに離れて存在していたり、関連するものが互いにつながっていたり、重複していたり、あるいは高等生物では一つの遺伝子がばらばらに分かれて（遺伝子でない塩基配列が介在して）存在したりと、その存在様式は多様である。

2. 3. 1 タンパク質について

タンパク質には生体内でいろいろな働きをする沢山の種類がある。例えば生体内化学反応を促進して物質やエネルギーの代謝に関与する酵素をはじめ、神経伝達物質やホルモンや誘引・忌避物質、さらには温度や光や音等の感覚を受け取る情報受容体、運動や物質輸送に関与するもの、免疫機構で働くもの、生体の構造を維持するものなど、あらゆる生命現象に関与し、生命を維持する必須の物質である。

タンパク質は20種類のアミノ酸が100~500個、ものによっては約1,000個も鎖状に結合している高分子である。それで遺伝情報の担い手である核酸を情報高分子というのに対して、タンパク質は機能高分子という。タンパク質もまたアミノ酸の1文字または3文字記号の配列として、文字列で示すことができる。

2. 3. 2 遺伝暗号について

特定のタンパク質をつくる遺伝情報とは、特定のアミノ酸配列をしたタンパク質をつくる情報と言い換えることができる。それでは遺伝情報はどのようにしてタンパク質のアミノ酸配列を規定しているのか。3個連続した塩基の組をコドンというが、このコドンはアミノ酸に対応し、コドンの配列がタンパク質のアミノ酸の配列を決めているのである。アミノ酸とコドンとの対応関係、つまりアミノ酸に対する暗号を遺伝暗号 (genetic code) という。タンパク質合成に直接関わるRNAコドンによる遺伝暗号表 (codon table) を表1に示す。3連続塩基の組は64通りあるが、このうち61通りはどれかのアミノ酸の暗号であり、3通りはタンパク質の合成の終了の信号である。メチオニンというアミノ酸の暗号は合成開始の信号を兼ねている。アミノ酸によって対応する暗号は1~6種類ある。遺伝暗号はほとんど全ての生物に共通であるので、普遍暗号 (universal code) と呼ばれる。

表1 遺伝暗号表
(コードとそれに対応するアミノ酸)

		第 二 塩 基				
		U	C	A	G	
第 一	U	UUU フェニルアラニン	UCU セリン	UAU チロシン	UGU システイン	U
		UUC フェニルアラニン	UCC セリン	UAC チロシン	UGC システイン	C
		UUA ロイシン	UCA セリン	UAA 終 結	UGA 終 結	A
		UUG ロイシン	UCG セリン	UAG 終 結	UGG トリプトファン	G 第
一	C	CUU ロイシン	CCU プロリン	CAU ヒスチジン	CGU アルギニン	U
		CUC ロイシン	CCC プロリン	CAC ヒスチジン	CGC アルギニン	C
		CUA ロイシン	CCA プロリン	CAA グルタミン	CGA アルギニン	A 三
		CUG ロイシン	CCG プロリン	CAG グルタミン	CGG アルギニン	G
塩 A	AUU イソロイシン	ACU トレオニン	AAU アスパラギン	AGU セ リ ン	U	
		AUC イソロイシン	ACC トレオニン	AAC アスパラギン	AGC セ リ ン	C 塩
		AUA イソロイシン	ACA トレオニン	AAA リ シ ン	AGA アルギニン	A
		AUG メチオニン	ACG トレオニン	AAG リ シ ン	AGG アルギニン	G
基 G	GUU バ リ ン	GUU アラニン	GAU アスパラギン	GGU グリシン	U 基	
		GUC バ リ ン	GCC アラニン	GAC アスパラギン	GGC グリシン	C
		GUA バ リ ン	GCA アラニン	GAA グルタミン	GGA グリシン	A
		GUG バ リ ン	GCG アラニン	GAG グルタミン	GGG グリシン	G

(注) タンパク質合成開始の暗号を兼ねる

3 遺伝情報系と言語情報系の比較

3. 1 遺伝情報の伝達と発現

前述のように、遺伝情報の二、三の性質は通常言語とよく似ていることがわかる。そこでさらに遺伝情報の伝達・発現を簡易化したスキームで見ると、それは通常言語における情報伝達・表現と同じパターンであることがより一層鮮明になる(図4)。細胞分裂の際にDNAは

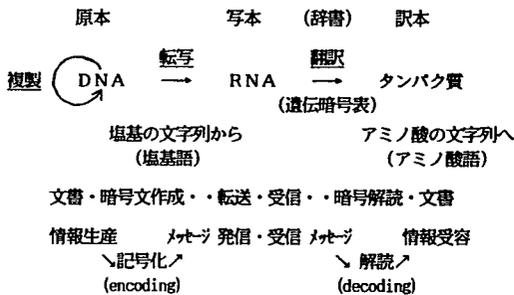


図4 情報の流れ

複製され、同じ遺伝情報が二つの新しい細胞に伝達される。情報が発現してタンパク質をつくる際には、まず遺伝子のもつ情報はRNAにコピー(転写)される。この情報コピーを使ってタンパク質がえられる。塩基の文字列から同じく塩基の文字列へと転写は起こるが、タンパク質をつくるときは、塩基の文字列からアミノ酸の文字列へと変換が起こる。つまり、塩基語からアミノ酸語への翻訳である。

この遺伝情報の発現の過程は、まさに原本、写本(副本)、訳本の関係に相当する。また文書から暗号文を作成して発信し、受信した暗号文を解読して文書を再生する過程と同じである。もっと一般化して言えば、情報生産、記号化、メッセージの送受信、受け手による解読、情報受容のスキームとなんら変わるところがない。遺伝情報系では送り手と受け手が人ではないという点で異論はあろうが、通常の言語系に対して遺伝言語系とすることが可能であろう。なお、遺伝情報の発現の英語は表現の英語と同じ

expressionであるが、分子遺伝学や生化学では発現を用語としている。

3. 2 言語は暗号化されている。

この音とこの音とをこの順序に並べると、こういう意味を示す、という自然発生の約束事(取り決め)から言語系は成り立っている。したがって音は暗号(記号)または暗号の構成要素であり、約束事にしたがって暗号は解読される。音声記号であろうと文字に変換した文字記号であろうとそれは同じである。我々は母国語を使うときにはあまり辞書のご厄介にはならない。それは、生まれたときから耳からはいて、脳にしっかり記憶されているからである。いわば母国語の辞書を脳の中に持っているようなものである。しかしあまり使わない言葉とか難解な言葉に出会うと、脳の外にある辞書をひく。外国語を読むときは母国語を読むときよりもよく辞書をひく。

翻訳のときに辞書を用いることから分かるように、文章の中には情報が隠れている。情報とはいわば潜在状態にある意味であるともいえる。情報は顕在化すると意味として表現され、意味を人に与える。遺伝情報は遺伝暗号を用いて翻訳される。したがって遺伝暗号表は辞書と同じと見なすことができる。遺伝情報もタンパク質として発現(顕在化)され、タンパク質は生体内で特定の意味のある働きをする。つまり、遺伝情報は生物に対して意味を与える。

3. 3 塩基の配列順序は重要である。

塩基の並ぶ順序には意味がある(遺伝暗号表を参照)。順序を入れ替えたなら別の意味になる。遺伝暗号が3文字だから、同じ3文字の通常言語と比較してみる(表2)。

表2 記号の配列は言語の構成の重要な一要素
アンダーラインは意味のある配列を示す

<u>AUG</u>	<u>UCG</u>	ワタシ(私)	コトバ(言葉)	<u>has</u>	<u>ten</u>	<u>cat</u>
<u>AGU</u>	<u>UGC</u>	ワシタ	コバト(小鳩)	hsa	tne	cta
<u>UGA</u>	<u>CGU</u>	タシワ	トバコ	<u>ash</u>	ent	atc
<u>UAG</u>	<u>CUG</u>	タワシ	トコバ	ahs	etn	<u>act</u>
<u>GAU</u>	<u>GUC</u>	シワタ	バコト	sha	nte	tca
<u>GUA</u>	<u>GCU</u>	シタワ	バトコ	sah	<u>net</u>	tac

塩基の配列を変えたものはすべて意味があり、違った意味になる場合もある。一方通常言語では、記号の入れ替えによって無意味なものが多くできる。このように言語の構成にとって記号の配列は一つの重要な要素であり、遺伝情報系でもまったく同様である。

通常の言語では、日本語49文字（音声記号に対応するカナ文字として）、英語26文字と多いが、遺伝情報系ではたった4種の文字を十分に活用している。どの3文字の組み合わせも意味を持っている。したがって遺伝暗号の辞書は僅か64語だけでよいが、通常言語の辞書は大きくて、しかもすべての語を含めることは困難である。コンピュータ言語でも語彙は多い。

この点からは、遺伝情報系の方が通常の言語情報系よりも言語として見ると優れているといえるであろう。この違いは、生命の起源以来の遺伝情報の歴史とそれより遥かに短い言語の歴史との差によるのかもしれない。

3. 4 遺伝情報と言語情報の階層的対比

何々の遺伝子という遺伝子の名前は、図書・論文などの題名 (title)、遺伝暗号を構成する3連続塩基 (コドン) は意味を持った単語、コドンが配列している遺伝子 (文字列) は文章に相当している。高等生物の細胞には複数の染色体があり、それぞれの染色体を構成するDNAは異なる遺伝子を担っている。ヒトでは約10万個の遺伝子が23対の染色体に分散して存在している。個々の染色体はいわば全集や百科事典の各巻に相当するといえるであろう。またこの遺伝子はこの時期に、その遺伝子は別の時期に、

というように多数の遺伝子が必要な時期に自動的に働いて、必要なタンパク質を必要な量だけつくる。その遺伝子の協調、言い換えれば遺伝情報の発現の調節ということがあって、複雑な生命現象が現れる、この情報の発現は巧妙に制御されていて、その制御機構は遺伝言語のきわめて精巧な文法に当たるといえるかもしれない。それに比べると、通常言語の文法は比較的簡単ということになるだろうか。しかし、この点に関しては今後の検討を要する。

3. 5 遺伝情報の言語としての位置づけ

前節において、遺伝情報と言語情報の構造を階層的に対比させて両者の類似性を指摘したが、遺伝情報系を一種の言語系と見なして、その構造的側面だけでなく機能面についても、他の言語系と若干の比較を試みたのが表3である。動物語を言語としてよいかどうかは見解の別れるところであるが、後述するように、ここでは広く言語に含める。

自然言語であろうと人工言語であろうと、その機能を端的に言えば、情報の伝達と表現 (発現) ということになる。エスペラント語は通常言語と変わらない。これはコンピュータ言語でも同じで、情報処理の中身は情報の保存、伝達、表現と言い換えられる。

言語の伝達の対象と言語を使用しようとする原動力になるものは、当然のことながら言語によって異なる。コンピュータ言語では人から機械 (物体) へであり、人から人ではないが言語と呼んでいるので、遺伝情報系では分子 (物質) から分子へであっても、遺伝言語系と呼んでもよさそうである。

通常言語の文章は、単語の意味の単なる集積とは違ってはるかに多量の意味をもつ。遺伝情報でも、文章に相当する遺伝子はコドン (個々の遺伝暗号の実体) のもつ意味の単なる集合ではない別の深い意味をもつ。この点も通常言語と共通している。これは人工言語でも同様であ

表3 遺伝情報の言語としての位置付け

言語の種類	使用の目的、理由	伝達 (動機)	意味をもつ最小単位	実 体	記 号
◎自然言語					
・日本語 ・英語 ・仏語など	情報の伝達と表現: [意思、要求、概念 思想、文化の提示、 伝達等]	人→人へ (意志、本能)	音素の組合わせと その連続＝単語	音声[音波、 電磁波の 一種] ①	文字 日本語49種 英語26種
・動物語 サル、 イルカなど	情報の伝達と表現: [警報、告知、要求、 求愛(種の保存)等]	動物→動物へ 動物→人へ (本能)	同上	同上	None
・遺伝語	情報の伝達、発現: [タンパク質の形成、 究極的には生命の積 極的維持に導く]	分子→分子へ (本能以前?)	コドン (遺伝暗号) ＝ アミノ酸を指定 する3連続塩基)	DNA の塩基 [分子]	文字 4種 A, G, C, T
◎人工言語					
・エスペラント	通常言語に同じ	通常言語に同じ	通常言語に同じ	通常言語に 同じ 磁気	文字
・BASIC ・FORTRAN ・COBOL ・C など	情報処理: [保存、加工編集、 印刷、伝播など]	人→機械へ (人の意志)	人に対しては自然言 語と同じ 機械に対しては磁気 パタンの組合わせ	磁気	文字 (通常言語 の)、数、記号 等: 磁気ボタン と対応

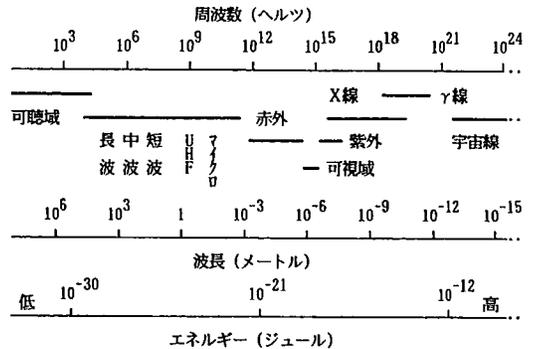
① 映像、画像の情報の実体は光 [光波 (電磁波の一種)] .

る。

通常言語を構成する音声はそれ自体記号であるが、この音声の実体は音波であり、電磁波の一種である。ちなみに映像や画像等の情報の実体は光波であり、また通常言語を文字記号で表した文章を読むときには光として感知する。これも音波とは周波数の領域が違うが電磁波の一種である。電磁波は粒子の性質をもち、それは波長の短い領域で著しく現れる。参考のために、電磁波のスペクトルとそのエネルギーの強さとの関係を周波数 (振動数) を座標にして表4に示した。

一方、遺伝情報を構成する実体は塩基で、物質である。物質の起源は宇宙の誕生にさかのぼる。現在の宇宙物理学によれば、超高密度に詰まったエネルギーの点が、約150億年前に突如爆発的に膨脹を開始し、その膨脹は現在もなお続いているが、そのごく初期にまずニュートリノが生成し、やがて陽子と電子が生じ、これが

表4 電磁波の周波数・波長・エネルギー



組み合わせさせて最も簡単な原子である水素原子が生成した。4原子の水素が融合すると、その質量の約0.7%をエネルギーとして放出してヘリウム原子が生じる。そのようにして原子核融合や核分裂を繰り返して多くの他の原子が生成し (原子進化), 原子間の化学結合で分子を生じ、分子は低分子から高分子へと進化して生体高分子ができる。この化学進化はやがて、約46

億年前に誕生した地球上で約10億年を経て生命の誕生にこぎつけることになる。このように見ると、質量はエネルギーと等価であるとしたA. Einsteinの理論をもち出すまでもなく、質量とエネルギーとは相互に変換する。したがって塩基という物質と電磁波との間に本質的な違いは無いことになる。

4. 考察

4. 1 動物語に関して

言語学者の多くは、動物は言語をもたないという。例えば、“遺伝言語と音声言語は類似した構造をもつが、……言うまでもないことだが、両者の間には何の因果関係もないことは、言語をもたない人間以外の動物でもその体質中に「遺伝言語」が含まれている点を考えれば明白である” [10]、と動物語をまったく度外視したうえで、遺伝情報と言語との関連性を否定している。そして人だけが言語をもち、それも数万年（さかのぼっても約20万年前）からという。しかしこれは言語をどう定義するかに依存する。例えば、言語は音声による意思の伝達の手段であるとすれば、言語をもつのは音声を発する動物に限定され、さらにある意図をもって話すことができるか、発声器官の構造が多様な音声を出せるようになっていたりとか、短期記憶力が十分であるとか、大脳左半球優位性が見られるとか、音素を組み合わせる能力があるとか、いろいろ条件を並べると人間しか言語をもてなくなる。しかしサバンナモンキーなどのように物事を明確に仲間に伝えるサルもいる [11]。ヒトとサルの違いを強調して定義するか、言語のごく基本的な性質のみで定義するかの問題である。確かに人の言語と動物の言語とは質的に違う。音素を組み合わせる単語をつくり、単語の組み合わせで節や文章をと、階層的な言語を形成するのは人間だけであろう。言語の階層性があるかないかが人の言語と動物言語の差であり、人の言語情報系と動物の情報

伝達系との間に進化的な関連がないとはいえない。この点、遺伝情報系は言語としての階層性がある。イルカ語を研究している人もいるし、何よりもイルカもサルも仲間同士でコミュニケーションしており、我々とはときにはイヌと心が通じ合えたと思うことさえある。それで表3は、動物の発する音声も動物語として自然言語に含めた。

4. 2 遺伝情報言語論の展開

表3では、言語の構造と機能の観点から、遺伝情報を言語と見なして自然言語に位置づけたが、これは比較の一部を示したに過ぎない。言語構造についてさらにいうと、それは主語と述語とからなる。また名詞と動詞や形容詞、副詞などの区別もある。遺伝子産物のタンパク質にもこれに類する構造が認められる。例えばウレアーゼという酵素タンパク質は尿素を加水分解してアンモニアと二酸化炭素とにする。その機能を果たすために、尿素を認識して結合する構造の部分（領域）と加水分解の作用をする領域とがある。いわば酵素には何をどのようにどうする、の部分それぞれあるということである。簡単な作用をする酵素は名詞と動詞からなり、複雑な作用をする酵素には他の品詞にあたる領域や複数の句からなるものがある。タンパク質の構造領域に対応した構造をその遺伝子もっている。このような例示は他にもできるが、それは生化学や分子生物学で扱う個々の事象の細部に踏み込むことになるのでこの程度に止める。

F. de Saussure は言語記号を2つの成分、signifiant（意味するもの；意味を与えるもの）と signifié（意味されるもの；意味そのもの）からなる分解できない統合的な実体と見なしているが [12]、signifiant は知覚し得るもので記号構造に、signifié は理解され得るもので記号の与える意味に相当すると考えると、塩基という記号で書かれた遺伝子構造は signifiant であり、遺伝子発現の産物であるタンパク質の機

能、すなわちその与える意味は signifié である
と対比できる。3. 2 で述べたように、DNA
が担っている情報とはいわば潜在状態にある意
味であり、情報はタンパク質として顕在化（発
現）して機能するのである。

分子遺伝学の分野では transcription,
translation, expression, encode, decode 等
の他に、transformation, transfer, transposition
等の用語もよく使用するが、これらは言語学の
用語でもあり、意味内容もよく似ている。これ
は単なる比喩として使われているだけであろう
か。また遺伝情報系はすべての生命現象の基盤
をなすシステムであるから、生命的なもののす
べての現象はこれに源を発しているとは考えら
れないだろうか。人間が発明したコンピュータ
は 0, 1 の 2 進法を用いているが、これは人の
脳における神経伝達の基本的メカニズムが、ニュー
ロン（神経細胞）に電気信号（神経インパルス）
を発生するかしないかという 2 進法であるのと
同じである。脳ではこのデジタル信号の集積
として神経伝達がアナログ的に行われる。人間
の思考や行動の範囲は、人間に本来備わってい
るもの、つまり人間を構成している構造やその
働きの範囲から、あるいは 1 章で触れたように
既体験の範囲から所詮は逸脱できないのではな
かろうかと思われる。もしかしたら言語につい
ても、遺伝情報系を背景にしながらそれを意識
することなく自然に形成されたかも知れない。
とするならば遺伝情報系と言語情報系の類似は
当然かもしれないのである。

微生物の遺伝生化学の研究から 1 遺伝子-1
酵素仮説を提唱（1945）して 1958 年にノーベル
医学生理学賞を受賞した George W. Beadle の
著書『The Language of Life』を、著名な言語
学者の Roman Jakobson が引用して、“（この
著書の題名は）単なる比喩ではなくて、遺伝情
報体系と言語情報体系との間の異常な程の類似
は、「象形文字よりもはるかに古く、生命その
ものと同じくらいに古い言語であり、しかもす
べての言語のうち最も生命のある言語を我々が

所有していることを、DNA コードの解明は明
らかにした」というこの著書の主題を十分に正
当化している”と述べている [13]。

細菌酵素の誘導合成の遺伝的制御の研究
（1961）により、1965 年にノーベル医学生理学
賞を受賞した F. Jacob と J. Monod は、遺伝
子の情報を転写してタンパク質合成の場に情報
を提供する物質をメッセンジャー RNA（遺伝
的メッセージを運ぶ伝令 RNA）と命名したが、
Jakobson は Jacob の講演を引用して、“すべ
ての情報伝達系のうちで、それ自身は意味をも
たないが、最小有意味的単位……の構成に用い
られる離散的な要素（音素とか塩基のこと）の
使用を基礎とする体系は、遺伝コードと言語コ
ードだけである”と主張している。また Jacob,
Jakobson, Lévi-Strauss らによれば、“ある一
定の仕方では並べられるとある意味をもつよう
な諸単位……（における）意味は、言語の場合
には単語の意味であり、遺伝の場合には……化
学的遺伝メッセージに沿って書き込まれた諸機
能を表すための意味である” [13]。Jakobson の
この見解は、彼の他の著書 [14] の Lévi-Strauss
による序文の中でも触れられている。さらに分
子生物学の分野の著名な研究者である F. H. C.
Crick や C. Yanofsky らの遺伝コードに関する
総説にも Jakobson は注目しており、彼はすで
に 1973 年頃には遺伝情報の言語性に強い関心
を抱いていたことは明らかである。

分子生物学者の Robert Pollack は最近の著
作 [15] の中で、ヒトの DNA は文学作品で偉
大な歴史物語であり、その作家は自然選択であ
ると考え、化学物質でつづられた物語という比
喩はイメージだけのものでなく、DNA の構成
分子（塩基のこと）は、機能的には本の中の文
字のように列となっていると述べている。また
彼は、生物学はここ数十年間の発展により、全
く別の学問、例えば記号論、と接点をもちはじ
めた、と認識して「DNA 言語」という言葉
を用いている。さらに彼は、言語に関わる遺伝子
は何万もあって、少なくとも脳の発語に関する

Brocaの領野において互いに絡み合っていて働いているに違いないとか、あるいは、ヒトが地域的に分離して起こる言語の分化と遺伝子の分化とを関係づける研究が新しい学問分野として発展するのではないかと記しているが、こうした見解は興味深いものがある。

最近、多くの生物のゲノムDNAの無意味な部分が言語を表現しているらしいことを示唆する研究が行われている [16]。この部分は junk DNA と呼ばれ、タンパク質に対する遺伝子でない非翻訳領域で、遺伝子の数十倍もあるが、長い進化の過程で蓄積した無意味な部分と考えられてきた。何種類かの言語の原文からいくつかの単語の出現頻度をランク付けして、これらの単語の頻度のランクを出現頻度に対してプロットすると、ある特定の関係が得られる。例えば、一番頻度の高い the は10番目に頻度が高い単語の約10倍の頻度で出現し、100番目の頻度の単語の100倍もよく使われるという。それで種々の生物のDNAについて、3から8個までの長さの塩基対を単語に見立てて、これらの配列の出現頻度とそのランクとの関係を調べると、この関係は junk DNA の部分にだけ現れた。すなわち junk DNA は言語の規則に従っているというのである。果たしてこの塩基配列の組織性は言語を支配する遺伝情報を示すのか、それとも生物の言語としての意味をもつのか、あるいは単に言語構造と似た特徴が見られたというだけなのか、今のところ謎である。

4. 3 言語の進化的側面

僅か4種類の文字からなる遺伝言語は通常自然言語と比べると、当然のことながら伝達の誤りや書き写しの誤りがきわめて少ないと考えられる。これは遺伝という現象を考えるとときに特に重要である。誤りが多くては人はすぐに人でなくなってしまうし、致死的でもある。これでは遺伝は成立しない。実際に誤りはごく少ないし、その上に伝達の誤りを修復する機能があ

るので、一層ミスは少ない。

それでは、文字が少ないから伝え得る情報は限られるかということそうではない。実に多くのことを伝えていて、それによってこの複雑で神秘的な生命現象のすべてを司っているのである。

我々は日常会話であるいは会議のときに、真意を伝えるのにときとして困難を感じる。互いに相手の言ったことを間違えて受け取って、自分の意見を述べたりする。「いや、そういう意味ではない」といって、より丁寧に発言を繰り返したり、言い換えたりしなくてはならないことがしばしばある。通常の言語は過去数10万年の間にどんどん進化し、現在でも変化し続けている。まだまだ幼稚で進化の余地が大いにあるといえるであろう。

遺伝言語ではそういう間違いは許されない。それは致命的になるからである。従って遺伝言語は基礎構造(遺伝暗号のレベル)においては保守的で、もはやあまり進化しないと考えられる。遺伝暗号は生物進化のごく初期の頃に進化し、その後の生物進化の過程においてほとんど変化しないで、ほとんどの生物に普遍的である。しかし遺伝暗号の配列、つまり遺伝子のレベルではまだまだ進化の余地は十分過ぎるほどある。たかだか100個ほどのコドンの並び方は無数といっていほどあり、生物はタンパク質を指定するのにそのうちのごく僅かの配列しか使っていない。

4. 4 今後の課題

以上の考察から、遺伝情報系は言語情報系と単に高度に類似しているというだけではなく、それは一種の言語系であると認めることが可能であろう。こうした考えをもっているのは著者だけではない。これを認めるには、言うまでもなく言語の定義を拡張することが前提になる。

それでは、遺伝情報系に言語情報系と同じ性格を与えている情報の本質は一体何であろうか。情報を人文・社会科学的に眺めただけでその本

質が把握できるであろうか。情報の本質を知るということは、情報を普遍的な視野で捉えるということである。それには冒頭で述べたように、自然科学からも眺めてみる必要があるだろう。本論文はその一環としての意味をもっているであろう。しかし自然から考察するとしてもその角度は非常に広い。生化学における個々の情報の問題だけでもきりが無い。人の生存する地球を含めて宇宙の起源にまで手を広げることは大変な作業である。しかし情報は人が生産し人に伝達するものであると、人にこだわる視野もあるので、今後は情報の生産と受容に働く脳、特に人の脳内で起こるできごとに焦点をあてた自然科学的考察が必要であろう。

謝辞

Jakobson の資料の収集にあたり、協力を得た愛知淑徳大学図書館四谷あさみ君に謝意を表す。

文献

- [01] 武山信子, 竹村彰祐. 生命における自己組織化 : 秩序の形成 (Self-organization of life : formation of ordered structure). *J. Libr. Inf. Sci.* Vol. 9, p.1-22 (1995)
- [02] Belkin, Nicholas J. and Robertson, Stephen E. Information science and the phenomenon of information. *J. Am. Soc. Inf. Sci.* Vol. 27, No 4, p. 197-204 (1976)
- [03] 情報学の理論と実際. 津田良成, 上田修一監訳. p.41-45. 東京, 勁草書房, 1995 ; (Vickery, Brian C. and Vickery, A. Information science in theory and practice (1987))
- [04] 竹村彰祐, 大井龍夫. 生物の情報システム. 東京, 講談社サイエンティフィック, 1990. 182p.
- [05] 渡辺智山. 価値観の生成過程に関わる問題点—知識の活性化と感情との関連を中心に— (Constructing the model of 'view of value' : focusing on the relations between knowledge activation and emotion). *Libr. Inf. Sci.* No 32, p. 141-162 (1994)
- [06] Watson, James D. and Crick, Francis H. C. A structure for deoxyribose nucleic acid. *Nature*, Vol. 171, p. 737-738 (1953)
- [07] 三浦謹一郎. 核酸の化学. 序 (江上不二夫による). 東京, 東京化学同人, 1962.
- [08] 大澤省三, 高木康敬, 竹村彰祐, 磯晃二郎, 関口睦夫, 核酸—その生物学・化学・物理学—. p. 85, p. 435, p. 464. 東京, 広川書店, 1963.
- [09] 分子生物学の夜明け (上). 野田春彦訳, p. 272, p. 13. 東京, 東京化学同人, 1982 ; (Judson, Horace F. The eighth day of creation (1979))
- [10] ロマーン・ヤコブソン選集 2, 言語と言語科学. 服部四郎編, 早田輝洋他訳. はしがき, vi. 東京, 大修館書店, 1978.
- [11] 正高信男他. 特集 言葉の起源. *日経サイエンス*, Vol. 24, No 10, p. 16-43 (1994)
- [12] ロマーン・ヤコブソン選集 2, 言語と言語科学. 服部四郎編, 早田輝洋他訳. 言語の本質の探究, p. 67. 東京, 大修館書店, 1978 ; (Jakobson, Roman. *Essais de linguistique générale II* (1973))
- [13] ロマーン・ヤコブソン選集 2, 言語と言語科学. 服部四郎編, 早田輝洋他訳. 言語の科学と他の諸科学との関係, III 言語学と自然科学, p. 195, p. 196. 東京, 大修館書店, 1978 ; (Jakobson, Roman. *Essais de linguistique générale II* (1973))
- [14] 音と意味についての六章. 花輪 光訳, 序 (レヴィーストロースによる), p. 12-13. 東京, みすず書房, 1977 ; (Jakobson,

Roman, Six leçons sur le son et le sens
(1976))

- [15] DNAとの対話：遺伝子たちが明かす人間社会の本質. 中村桂子他訳. p. 12-13, p. 20-21, p. 90, p. 221, p. 224-228. 東京, 早川書房, 1995 ; (Pollack, Robert. Signs of life: the language and meanings of DNA (1994))
- [16] Yam, Philip. ジャンクDNAに現れた言語. 日経サイエンス (Overseas コラム), Vol. 25, No. 5, p. 104-105 (1995)