

バイオインフォマティクスシリーズ 1

# バイオインフォマティクスのための 生命科学入門

浜田 道昭 監修  
福永 津嵩 共著  
岩切 淳一

コロナ社

## シリーズ刊行のことば

現在の生命科学においては、シーケンサーや質量分析器に代表される計測機器の急速な進歩により、ゲノム、トランスクリプトーム、エピゲノム、プロテオーム、インタラクトーム、メタボロームなどの多種多様・大規模な分子レベルの「情報」が蓄積しています。これらの情報は生物ビッグデータ（あるいはオミクスデータ）と呼ばれ、このようなデータからいかにして新しい生命科学の発見をしていくかが非常に重要となっています。

このような状況の中でその重要性を増しているのが、生命科学と情報科学を融合した学際分野である「バイオインフォマティクス」（生命情報科学，生物情報科学）です。バイオインフォマティクスは、DNA やタンパク質の配列などの、生物の配列情報をデジタル情報として捉え、コンピュータにより解析を行うことを目的として誕生しました。このような、生物の配列情報を解析するバイオインフォマティクスの一分野は「配列解析」と呼ばれます（これは本シリーズでも主要なテーマとなっています）。上述の計測機器の進歩とともに、バイオインフォマティクスはここ数十年で飛躍的に発展し、いまや配列解析にとどまらずに、トランスクリプトーム解析，メタボローム解析，プロテオーム解析，生物ネットワーク解析など多岐にわたってきています。また，必要な知識も，統計学，機械学習，物理学，化学，数学などの多くの分野にまたがっています。しかしながら，これらのバイオインフォマティクスの多岐にわたる分野を，教科書的・体系的に学ぶことができる成書シリーズは，国内外を見てもほとんどありません。

そこで，大学生，大学院生，技術者，研究者などに，バイオインフォマティクスの各分野を体系的に学習することを可能とするための教科書を提供することを目的として本シリーズを企画しました。これを実現するために，バイオイン

フォマティクス分野の最前線で活躍をしている、若手・中堅の研究者に執筆を依頼しております。執筆者の方々には、バイオインフォマティクス研究の基盤となる理論やアルゴリズムを中心に、可能な限り厳密かつ自己完結的に解説を行うようお願いしています。そのため、本シリーズは、大学などにおけるバイオインフォマティクスの講義の教科書として活用可能であるのみならず、読者が独学する場合にも最適な書籍になっていると確信しています。

最後になりますが、本シリーズの企画の段階から辛抱強くサポートして下さったコロナ社の皆様に御礼を申し上げます。本シリーズが、今後のバイオインフォマティクス研究さらには生命科学研究の一助となることを切に願います。

2021年9月

「バイオインフォマティクスシリーズ」監修者 浜田道昭

# ま え が き

ゲノムデータをはじめとする生命科学ビッグデータが日常的に産出される現在、その情報解析を担うバイオインフォマティクスはいまや生命科学にとって必要不可欠な存在である。そのため、機械学習やアルゴリズムといった数理情報学的知識を持った技術者・研究者が、生命科学において今後いっそう重要な役割を果たしていくことは間違いない。ただし、ビッグデータ分析においてより適切なモデリングや解析を行うためには、対象となるデータ、つまり生命科学のデータについてある程度のドメイン知識を持っていることが重要となってくる。しかしながら、生命科学の分野は非常に広大であり標準的な教科書は頁数が多く、また、数理情報学分野の研究者にはなじみのない暗記的作業がどうしても必要となる。このため、数理情報学の研究者が生命科学の勉強をすることは容易なことではなく、この点が生命科学への参入障壁となっているといえるだろう。

本書は「バイオインフォマティクスシリーズ」の第1巻として、数理情報学分野の学生・技術者・研究者を対象に、バイオインフォマティクスを行うために必要な生命科学の基礎について解説した教科書である。生命科学になじみのない読者は、本シリーズのほかの巻を読む前に本書を一読すると、ほかの巻が扱っている生命科学的な内容について理解が深まるだろう。なお、本シリーズのほかの巻が、バイオインフォマティクスのさまざまな問題の数理的背景を厳密に記述することを念頭に置いて執筆されているのに対し、本書には数式は一切出でこず、あくまで生命科学の教科書であるという違いがある。ただし、解説事項は生命科学データ解析において特に頻出する事項に絞って解説しており、生命科学になじみのない読者でもなるべく読みやすくコンパクトに必要な事項を理解できるように試みた。そのため、いくつかの項目では最先端の研究内容も

含まれているが、逆に標準的な生命科学の教科書には当然記載されているような内容にもかかわらず、本書にはまったく記述がないというような事項も存在する。そのため本書では、発生学や神経科学など、生命科学における多くの重要な研究分野を紹介することはできなかった。また解析するデータによっては、本書では紹介しきれなかった内容についてもその分野の知識を必要とする可能性が十分にある。その際にはもちろん各分野の勉強をしなければならないが、本書の内容を十分に理解することは、より進んだ勉強をする上で大きな助けとなることだろう。

本書は三つの章から構成されている。第1章では、生命現象を分子的な観点から捉える分子生物学のセントラルドグマについて解説する。特に、近年急速に進歩した技術である遺伝子配列決定技術を中心に、遺伝子配列ビッグデータがセントラルドグマの過程の解明に果たす役割を紹介する。第2章では、第1章で登場する生体分子が、実際に物理的にどのような構造をとるのか、そしてその構造がどのように生体的な機能と関係するのかについて、生物物理学的な観点を紹介する。第3章では、遺伝子配列決定技術の急速な進歩を受けて、分野が変容しつつある進化遺伝学と微生物学について、その基礎とバイオインフォマティクスの果たす役割について解説する。

なお、大量の配列・数値データを扱うバイオインフォマティクスでは、生命現象を支えているそれぞれの生体分子がどのような大きさ・形をしているかを意識する機会が少ない。そのため、本書では生体分子の立体構造の図をできるだけ多く盛り込んだ。図説に出てくる PDBID は、生体分子の立体構造データベースである Protein Data Bank (PDB) の ID である。生体分子の立体構造に興味を持った読者には、自分で構造データを取得し、実際の構造を眺めてみることをお勧めしたい。また本書では、重要な科学的発見を行った研究者の氏名を多く紹介し、特にその研究者がノーベル賞を受賞されている場合、そのことも併せて紹介することとした。これは、本書で紹介している事象が非常に重要な知見であるとみなされていることを伝えるとともに、科学的知見の背後にはそれを発見した研究者が存在すること（すなわち、研究とは人間が行う行為

であること), そしてその流れのなかに自身の研究活動が連なることを認識することが重要であると考えたためである。

本書を出版するにあたり多くの方々にご協力を賜った。本シリーズの監修者である早稲田大学の浜田道昭 教授には、構成の段階から数多くのコメントをいただいた。関西学院大学の藤博幸 教授には、草稿段階における多くの誤記や不正確な点をご指摘いただいた。また、東京大学大学院生の山内駿 氏と今野直輝 氏にも、原稿の不十分な点を数多くご指摘いただいた。分子の構造や相互作用については、量子科学技術研究開発機構の桜庭俊 博士にアドバイスをいただいた。コロナ社には、著者の原稿執筆の遅れにも寛大にご配慮いただき、また執筆に関して多くの助言をいただいた。ご協力いただいた方々に心よりの感謝を申し述べさせていただきたい。本書を読んだことで、生命科学データ解析に挑戦する数理工学分野の技術者・研究者が一人でも増えるならば、著者の望外の喜びとするところである。

2022年6月

福永津嵩  
岩切淳一

# 目 次

## 1. 分子生物学のセントラルドグマとオミクスデータ

1.1 セントラルドグマの基礎	1
1.1.1 DNA と 複 製	2
1.1.2 RNA と 転 写	6
1.1.3 タンパク質と翻訳	10
1.1.4 セントラルドグマ	15
1.2 オミクスデータの測定技術	17
1.2.1 DNA 操作の一般的実験手法	17
1.2.2 サ ン ガ ー 法	20
1.2.3 第二世代塩基配列決定法	23
1.2.4 第三世代塩基配列決定法	27
1.2.5 RNA の配列決定と発現量推定	30
1.2.6 質 量 分 析 法	34
1.3 ゲノムと遺伝子	38
1.3.1 生体分子としてのゲノム	38
1.3.2 配列データとしてのゲノム	40
1.3.3 ゲノム配列の多様性と参照ゲノム配列	42
1.3.4 遺伝子とアノテーション	44
1.3.5 ゲノムサイズと遺伝子数	46
1.3.6 ゲノム配列決定	48
1.3.7 ゲ ノ ム 編 集	50
1.4 エピゲノムと転写制御	55

1.4.1	転写因子	56
1.4.2	クロマチン構造	60
1.4.3	ヒストン修飾とヒストンバリエント	63
1.4.4	DNAメチル化	66
1.4.5	ゲノムインプリンティングとX染色体不活化	69
1.5	トランスクリプトームと転写後発現制御	72
1.5.1	転写産物の種類	72
1.5.2	選択的スプライシング	75
1.5.3	新規転写産物の発見	77
1.5.4	RNAの発現制御	77
1.5.5	RNA修飾とエピトランスクリプトーム	79
1.6	プロテオームとタンパク質機能	82
1.6.1	タンパク質の折り畳みと機能	82
1.6.2	細胞内局在	84
1.6.3	翻訳後修飾	86
1.6.4	タンパク質の分類	87
1.7	データベースとオミクスデータ解析	92
1.7.1	生命科学におけるデータベース	92
1.7.2	ゲノムブラウザ—複数オミクスデータの可視化—	94
1.7.3	生物ネットワーク解析	97
1.7.4	遺伝子オントロジー解析	101

## 2. 生体分子の高次構造と分子間相互作用

2.1	生体分子の立体構造決定法と構造データ	104
2.2	生体分子に働く力・化学結合	106
2.2.1	静電相互作用	107
2.2.2	水素結合	108
2.2.3	疎水性相互作用	109
2.2.4	スタッキング相互作用	111
2.2.5	ファンデルワールス力	112



2.3 生体分子の高次構造 .....	112
2.3.1 DNA の 構 造	113
2.3.2 RNA の 構 造	115
2.3.3 タンパク質の構造	117
2.4 分子間相互作用 .....	121
2.4.1 相互作用の特異性	121
2.4.2 DNA-タンパク質相互作用	122
2.4.3 RNA-タンパク質相互作用	124
2.4.4 RNA-RNA 相互作用	127
2.4.5 タンパク質-タンパク質相互作用	128
2.4.6 化合物-タンパク質相互作用	132
<b>3. 進化遺伝学・微生物学のためのバイオインフォマティクス</b>	
3.1 進化遺伝学のバイオインフォマティクス .....	134
3.1.1 進化遺伝学の基礎理論	135
3.1.2 進化遺伝学研究におけるバイオインフォマティクスの重要性	140
3.1.3 ゲノムに起こる小規模な変異	141
3.1.4 ゲノムに生じる大規模な変異	144
3.1.5 系統関係の表現法と系統分類学	149
3.2 微生物学のバイオインフォマティクス .....	153
3.2.1 微生物学研究の重要性	154
3.2.2 微生物学研究におけるゲノム情報解析の重要性	157
3.2.3 微生物の系統分類	159
3.2.4 微生物の構造的・生理的特徴	161
3.2.5 ウ イ ル ス	163
3.2.6 原核生物ゲノムの特徴とゲノムアノテーション	164
3.2.7 微生物の比較ゲノム解析	168

3.2.8 微生物の比較ゲノム解析の研究例 — プロテオロドプシン保有微生物の  
生存戦略 — 171

3.2.9 メタゲノム解析 173

付 録 ..... 177

引用・参考文献 ..... 180

索 引 ..... 188

# 1 分子生物学のセントラルドグマとオミクスデータ

bioinformatics

生命現象の多くは、人間がまだその仕組みを理解できていないきわめて複雑な機構であるが、これらの仕組みはいずれも、タンパク質をはじめとする生体分子を基本要素として構成されている。研究対象とする生命現象にどのような分子がどのように関わっているのかを明らかにする、すなわち分子的な観点から生命を理解しようとする学問は分子生物学 (molecular biology) と呼ばれ、DNA を中心とする遺伝の仕組みが理解された 20 世紀後半以降は生命科学の中心的な学問分野となった。また DNA 配列を決定する技術の革命的な進歩により、現在、DNA 配列データが爆発的に増大しており、その DNA 配列ビッグデータを情報科学的に解析するためのバイオインフォマティクス (bioinformatics) は、もはや生命科学にとって必要不可欠な存在となっている。本章では、分子生物学において中心的な概念であるセントラルドグマを対象に、その生命科学的な機構およびバイオインフォマティクスの果たす役割について紹介する。

## 1.1 セントラルドグマの基礎

生物は、われわれヒトのように複雑な生物から、細菌、出芽酵母、原生動物のような比較的単純な生物まで、膜 (脂質二重膜) に囲まれた細胞 (cell) と呼ばれる構造を基本単位として構成されている。なお、ヒトのように多数の細胞から構成される生物は多細胞生物と呼ばれ、細菌、出芽酵母、原生動物などの単一の細胞からなる生物は単細胞生物と呼ばれる。ヒトと細菌のように異なる生物間では、個体を構成している細胞は大きく異なっている。さらに、ほとん

どの多細胞生物でも、同一個体内にもさまざまな組織・器官があり、それぞれが異なる特徴を持った細胞で構成されている。例えばヒトの場合、脳組織を構成する神経細胞と、筋組織を構成する筋細胞とではまったく役割が異なっている。これらの細胞のなかには、分裂して自身と同じ細胞の複製をつくって増殖する、という機能を持つ細胞が存在し、そのような細胞はすべての生物に存在している。

これら細胞を構成しているのは、おおよそ70%が水であり、残りの30%の多くは、タンパク質 (protein) (15%) や **RNA** (ribonucleic acid) (6%), **DNA** (deoxyribonucleic acid) (1%), 脂質 (3%), 糖質 (3%) などの生体高分子が占めている。細胞が自身と同じ複製をつくって増殖する<sup>†</sup>、という機能を実現するためには、これら生体高分子の存在が必要不可欠である。本節では、これら生体高分子のなかでも、DNA, RNA, およびタンパク質の3種類の分子に焦点を当て、これら3種類の分子がどのような役割を担い、どのようにつくられるのかを概説する。特に、遺伝物質として同一のDNA分子のコピーをつくる“複製”, DNAの持つ情報に基づいてさまざまなRNAやタンパク質がつくられる際の“転写”, および“翻訳”という三つの過程を中心に解説する。

### 1.1.1 DNA と複製

細胞が分裂して増殖する際、分裂後の細胞（娘細胞）は、分裂前の細胞（母細胞）と同じ形や性質（形質）を受け継いでいる。また、個体が交配することで子が生まれた場合、その子の個体は親個体とよく似た性質を持つ。このように親の形質が子に受け継がれる仕組みは、DNAと呼ばれる**核酸** (nucleic acid) の一種が重要な役割を担っており、特に生殖が伴う場合にはその仕組みは**遺伝** (heredity) と呼ばれる。DNAは、リン酸と五炭糖のデオキシリボース、そして**塩基** (base) の三つが結合したヌクレオチド (nucleotide) を基本単位とし、複数のヌクレオチドが1本の糸のように連なった生体高分子である (図 1.1)。

<sup>†</sup> 1.4.1 項で紹介するように、細胞分裂においては必ずしも細胞が自身と同じ複製をつくるとは限らない（細胞分化やiPS細胞の事例など）が、ここでは話を単純化している。

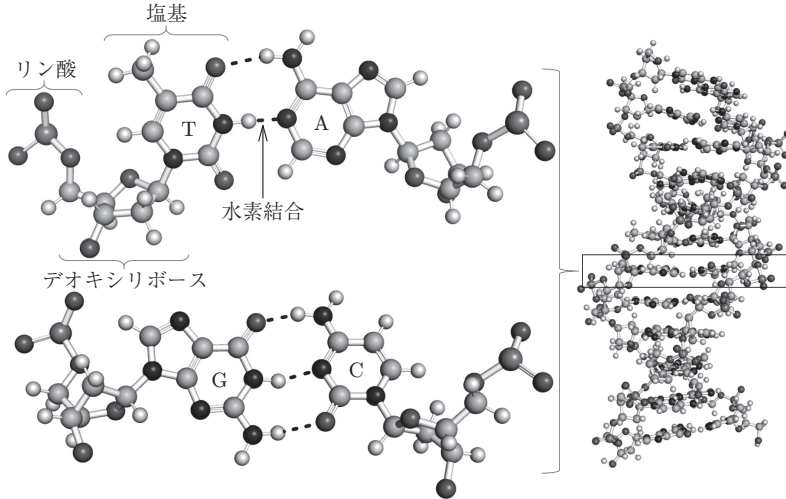


図 1.1 DNA の 4 種類のヌクレオチドが形成する塩基対と二本鎖 DNA の二重らせん構造 (PDBID ; 1D29)

DNA は 4 種類のヌクレオチドから構成されており、4 種類すべてにおいてリン酸と五炭糖が共通している。またヌクレオチド同士は、安定な共有結合 (covalent bond) によって 1 本の鎖状につながっている (図 1.2 ; この図では例のためヌクレオチドが二つしか連結していないが、現実には非常に多くのヌクレオチドが連結している)。このとき、結合には方向性が存在し、五炭糖における炭素の位置に基づいて 5' 側および 3' 側と呼ばれる。また DNA 鎖の両端

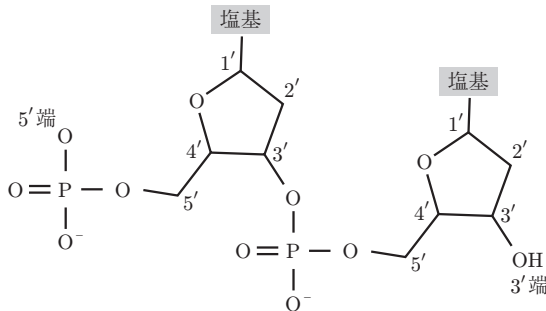


図 1.2 デオキシリボヌクレオチドの連結構造

は、その結合の方向性に基づいて、それぞれ 5' 端および 3' 端と呼ばれる。

なお、塩基には A (アデニン)、C (シトシン)、G (グアニン)、T (チミン) の 4 種類が存在する。そして、長い 1 本の DNA 分子のなかの 4 種類の塩基の出現順序を文字列として表現したものが DNA の塩基配列である (文字列の順序は 5' → 3' となっている)。すべての生物の細胞のなかには、遺伝物質として DNA が存在しているが、その DNA の保持の方式は細菌などの原核生物 (prokaryote) と動物や植物などの真核生物 (eukaryote) とでは異なっている。真核生物の細胞は真核細胞と呼ばれ、膜 (核膜) で隔てられた区画である核 (nucleus) にほとんどの DNA が存在している<sup>†</sup>。また、原核生物の細胞は原核細胞と呼ばれ、核のような膜で隔てられた区画がなく DNA やさまざまな分子が混在した状態で存在している。

20 世紀になり、遺伝物質の正体が DNA であることが明らかになって以降、DNA による遺伝の仕組みを明らかにするために、多くの研究者が DNA の 3 次元構造 (立体構造) を調べる研究の激しい競争を繰り広げた。DNA の立体構造解明のきっかけとなったのは、DNA 中の A と T の含有量が等しく、同様に C と G の含有量も等しいことがシャルガフ (Erwin Chargaff) によって発見されたことである。この性質は、シャルガフの法則と呼ばれている。さらに、フランクリン (Rosalind Franklin) によって、DNA がらせん状の構造であることが明らかにされた。その後、「シャルガフの法則を満たすようならせん構造」が 4 種類のヌクレオチドによってどのようにして実現されているのかを、1953 年にワトソン (James Watson) とクリック (Francis Crick) が分子模型を使った試行錯誤の末明らかにした。その DNA の立体構造とは、2 本の撚糸<sup>より</sup>のような DNA の間で、A と T、C と G がたがいにペアとして結合することで形成される二重らせん (double helix) 構造であった (図 1.1)。A と T は 2 本の水素結合 (hydrogen bond)、C と G は 3 本の水素結合でペアを形成しており、これらは塩基対 (base-pair) またはワトソン-クリック塩基対と呼ばれる。このと

---

<sup>†</sup> 真核細胞の内部構造については、1.6.2 項で詳しく説明する。

き、1本のDNAの塩基の並びが決まると、AとT、CとGの塩基対の法則に従って、もう1本のDNAの塩基の並びも自動的に決まることになる。この性質は、DNAの塩基配列の相補性 (complementarity) と呼ばれ、DNA分子の複製 (replication) を実現する基本原理として、すべての生物に共通している。2本のDNAが塩基対により結合している状態は、二本鎖DNA (double-stranded DNA)、1本ずつに分かれている状態は、一本鎖DNA (single-stranded DNA) と呼ばれる。なお、ワトソンとクリック、そしてフランクリンの同僚であったウィルキンス (Maurice Wilkins) の3人は1962年にノーベル生理学・医学賞を受賞している<sup>†</sup>。

細胞が二つに分裂する際、遺伝物質であるDNAが正確に複製され、二つの細胞にそのDNAが一つずつ分配されることで、母細胞の持つ遺伝情報が娘細胞に伝わる。この仕組みは、DNAの相補性と二重らせん構造に基づいて、つぎのように説明される。細胞分裂時には、まず一つの細胞のなかに存在している一つの二本鎖DNAが二つの一本鎖DNAに分かれる。そして、それぞれの一本鎖DNAを鋳型<sup>いがた</sup>として、相補的なDNA鎖がそれぞれ1本ずつ新しく合成されることで、二つの二本鎖DNAとして複製される。この二つの二本鎖DNAは塩基配列としてはまったく同じものとなっており、そしてこの二つの二本鎖DNAが娘細胞にそれぞれ一つずつ分配される。この分配された二本鎖DNAは、片方の鎖が母細胞にもともとあったDNAに由来しており、その相補鎖となるもう片方の鎖は新しく合成されたものである。このようなDNA複製の形式は半保存的複製と呼ばれている。

DNAの複製は、DNAが単独で自律的に行うものではなく、鋳型となる一本鎖DNAにさまざまなタンパク質が結合し、DNAの塩基対相補性の法則に従ってヌクレオチドを一つずつつなげることで、もう1本の相補的なDNAを合成

---

<sup>†</sup> フランクリンは1958年に死去したため、ノーベル賞受賞はならなかった (ノーベル賞は故人には授与されない)。またDNA二重らせん構造の解明は、フランクリンが取得した構造解明のための決定的なデータを、非正規な方法でワトソンとクリックが閲覧した上での解明であり、倫理的に大きな問題となった。

# 索引

## 【あ】

アクセサリーゲノム	170
アセチル化	64
アノテーション	46
アポ状態	132
アミノ酸	10
アレル	138
アンチセンス鎖	8
アンフィンゼンのドグマ	83

## 【い】

異質倍数化	146
一本鎖 DNA	5
遺 伝	2
遺伝暗号表	12
遺伝型	138
遺伝子	38
遺伝子オントロジー	101
遺伝子間領域	47
遺伝子水平伝播	148
遺伝子制御ネットワーク	98
遺伝子ノックアウト	50
遺伝子ノックダウン	50
遺伝子発現制御機構	56
遺伝的浮動	139
イントロン	45
インプリンティング	
制御領域	69

## 【う】

ウイルス	163
------	-----

## 【え】

液-液相分離	84
--------	----

エキソン	45
枝 長	149
エドマン分解法	34
エピゲノム	64
エピジェネティクス	64
エピトランスクリプトーム	
	82
塩 基	2
塩基対	4
エンハンサー	57

## 【お】

オオノログ	146
オーソログ	145
オフターゲット効果	53
オペロン	167
オミクスデータ	17
オングストローム (Å)	106

## 【か】

外 群	150
開始コドン	12
鍵と鍵穴モデル	132
核	4
核 酸	2
核磁気共鳴法	104
核タンパク質	89
化合物-タンパク質	
相互作用	132
カバレッジ	23
関 係	92

## 【き】

基 質	90
逆 位	144

逆転写	16
逆平行 $\beta$ シート	118
球状タンパク質	88
共有結合	3
極性分子	108
近接ライゲーション法	128
金属タンパク質	90

## 【く】

クロマチン	60
クロマチンリモデリング	61
クーロン力	107

## 【け】

形状相補性	130
系統ゲノミクス	151
系統樹	149
系統ネットワーク	151
系統プロファイル法	169
欠 失	141, 144
ゲノム	38
ゲノムアイランド	148
ゲノムアセンブリ	49
ゲノムインプリンティング	
	69
ゲノムブラウザ	95
ゲノム編集	51
ゲノムワイド関連解析	43
原核生物	4
顕 性	137

## 【こ】

コアゲノム	169
コイル	119
合成時解読法	24



酵素	90	ステム構造	115	短反復配列	142
構造タンパク質	91	スプライシング	45		
構造モチーフ	119	スプライソソーム	73	<b>【ち】</b>	
高速液体クロマトグラ フィー法	36			中立	139
酵母ツーハイブリッド法	130	<b>【せ】</b>		超個体	155
コード領域	45	生殖細胞	69	長鎖非コード RNA	74
コード領域密度	165	生成物	90	超二次構造	119
古細菌	159	静電遮蔽効果	107	重複	144
コドン	12	静電相互作用	107		
コドン表	12	正の選択	137	<b>【て】</b>	
<b>【さ】</b>		生物ネットワーク解析	98	低温電子顕微鏡法	105
細胞	1	ゼロモード導波路法	27	デオキシリボ核タンパク質	89
細胞内局在	84	繊維状タンパク質	88	適応度	136
細胞内小器官	84	全ゲノム重複	146	データベース	92
サンガー法	20	センス鎖	8	データベース管理システム	92
散在反復配列	147	潜性	137	デプス	23
参照ゲノム配列	43	選択的スプライシング	75	テロメア	39
<b>【し】</b>		セントラルドグマ	15	転移 RNA	14
シグナル配列	85	<b>【そ】</b>		電気陰性度	108
システムバイオロジー	101	相互作用の特異性	121	電気泳動法	17
次世代塩基配列決定法	23	操作的分類単位	160	転座	144
自然選択	136	相同	145	電子顕微鏡法	104
質量分析法	34	挿入	141	電子スプレーイオン化法	37
シャペロン	83	挿入欠失	141	転写	8
シャルガフの法則	4	相補性	5	転写因子	56
終止コドン	12	相補的 DNA	20	転写後制御	56
修復	141	側系統群	151	転写産物	8
主溝	114	疎水性	109	転写制御	56
シュードノット構造	115	疎水性相互作用	110	天然変性タンパク質	83
受容体タンパク質	91	疎水性プロット	111	天然変性領域	83
進化	135	<b>【た】</b>		点変異	141
真核生物	4	第三世代塩基配列決定法	27	<b>【と】</b>	
真菌	159	第二世代塩基配列決定法	23	同義変異	143
親水性	109	多系統群	152	糖タンパク質	89
真正細菌	159	多量体	128	トポロジカル関連	60
<b>【す】</b>		単系統群	151	ドメイン	119
水素結合	4, 108	単純タンパク質	89	ドメイン	119
スケールフリー	100	タンパク質	2	トランスオミクス	99
スタッキング相互作用	111	タンパク質コード遺伝子	45	トランスクリプトーム	72
		タンパク質-タンパク質		トランスポゾン	146
		相互作用	128		

		複合タンパク質	89			<b>【み】</b>	
		複製	5			ミスセンス変異	143
内 群	150	複製開始点	166			<b>【む】</b>	
ナンセンス変異	143	負の選択	137			無根系統樹	149
		プライマー	6			<b>【め】</b>	
		プラスミド	164			メタゲノム解析	158
<b>【に】</b>		フレームシフト変異	144			メタボローム解析	34
二重らせん構造	4	プロテオーム解析	34			メチル化	64
二本鎖 DNA	5	プロモーター	9			メチル化変動領域	68
		分子間相互作用	121			<b>【ゆ】</b>	
<b>【ぬ】</b>		分子進化の中立説	139			有根系統樹	149
ヌクレオソーム	39	分子生物学	1			誘導適合モデル	132
ヌクレオチド	2	分泌タンパク質	86			ユークロマチン	61
						ユビキチン化	87
						ゆらぎ塩基対	8
<b>【は】</b>						<b>【よ】</b>	
バイオインフォマティクス		<b>【へ】</b>				四重極型質量分析計	37
	1	平行 $\beta$ シート	118			<b>【り】</b>	
バイオレメディエーション	156	ヘテロオリゴマー	129			リード	22
		ヘテロクロマチン	61			リボ核タンパク質	89
バイサルファイト		ヘテロ接合	138			リボソーム	14
シーケンス法	68	ペプチドマスフィンガー				リボソーム RNA	14
配列モチーフ	59	プリンティング法	35			リポタンパク質	89
発 現	15	変 異	42			リボスクレアーゼ	77
パラログ	145					リン酸化	86
パンゲノム	43	<b>【ほ】</b>				<b>【る】</b>	
反復配列	47	補因子	90			ル カ	134
		芳香族アミノ酸	111			ループ	119
<b>【ひ】</b>		ホモオリゴマー	128			ループ構造	115
非共有結合	106	ホモ接合	138			<b>【れ】</b>	
非極性	109	ホモログ	145			レトロトランスポゾン	147
飛行時間型質量分析計	37	ポリメラーゼ連鎖反応法	18			<b>【ろ】</b>	
非コード RNA	10	ホロ状態	133			ロゼッタストーン法	120
ヒストン	39	翻 訳	14				
ヒストン暗号	65	翻訳後修飾	86				
微生物	153	<b>【ま】</b>					
微生物ダークマター	158	マイクロサテライト	142				
非同義変異	143	マイクロ RNA	74				
比誘電率	107	膜タンパク質	110				
表現型	138	マッピング	59				
		マトリックス支援レーザー					
<b>【ふ】</b>		脱離イオン化法	36				
ファンデルワールス力	112	マルチドメイン	119				
副 溝	114						

	<b>[A]</b>		<b>[L]</b>		
ATAC-seq 法	62	lncRNA	74	RNA-タンパク質複合体	89, 125
A-to-I 編集	81	LUCA	134	RNA-RNA 相互作用	127
				RNA-seq 法	31
				rRNA	14
				<b>[S]</b>	
<b>[C]</b>		<b>[M]</b>		SAG	159
CAGE 法	97	MAG	158	SBDD	133
ceRNA	78	miRNA	74	scRNA-seq	33
ChIA-PET 法	62	mRNA	9	SNP	42
ChIP-seq 法	58	mRNA 前駆体	45	SNV	42
CLIP-seq 法	126	MS	34		
CPR 群	174			<b>[T]</b>	
CRISPR 活性化	54	<b>[N]</b>		tRNA	14
CRISPR 干渉	54	ncRNA	10		
CRISPR-Cas9	51	NGS	23	<b>[X]</b>	
				X 線結晶解析	104
<b>[D]</b>		<b>[O]</b>		X 染色体不活化	70
DMR	68	OTU	160	~~~~~	
DNA	2			<b>【ギリシャ文字】</b>	
DNA オリガミ	114	<b>[P]</b>		$\alpha$ ヘリックス	117
DNA 結合タンパク質	122	PCR 法	18	$\beta$ シート	117
DNA ポリメラーゼ	6	PMF 法	35	$\sigma$ 因子	167
DNA メチル化	66	pre-mRNA	45	<b>【数字】</b>	
DNA-タンパク質相互作用	122			1 塩基多型	42
DNA-タンパク質複合体	89	<b>[R]</b>		1 塩基バリエント	42
		RIP-seq 法	126	1 細胞 RNA-seq 法	33
<b>[E]</b>		RNA	2	3C 法	61
EC 番号	90	RNA 結合タンパク質	125	3' 非翻訳領域	45
		RNA 構造プロービング法	117	5' 非翻訳領域	45
<b>[G]</b>		RNA 編集	81		
GWAS	43	RNA ポリメラーゼ	8		
		RNA-タンパク質相互作用	125		
<b>[H]</b>					
Hi-C 法	61				

—— 監修者・著者略歴 ——

浜田 道昭 (はまだ みちあき)

- 2000年 東北大学理学部数学科卒業
- 2002年 東北大学大学院理学研究科修士課程修了 (数学専攻)
- 2002年 株式会社富士総合研究所研究員
- 2009年 東京工業大学大学院総合理工学研究科博士後期課程 (社会人博士) 修了 (知能システム科学専攻), 博士 (理学)
- 2010年 東京大学特任准教授
- 2014年 早稲田大学准教授
- 2018年 早稲田大学教授
- 現在に至る

福永 津嵩 (ふくなが つかさ)

- 2011年 東京大学理学部生物情報科学科卒業
- 2013年 東京大学大学院新領域創成科学研究科修士課程修了 (情報生命科学専攻)
- 2016年 東京大学大学院新領域創成科学研究科博士後期課程修了 (メディカル情報生命専攻), 博士 (科学)
- 2016年 日本学術振興会特別研究員 (PD)
- 2017年 東京大学助教
- 2021年 早稲田大学講師
- 現在に至る

岩切 淳一 (いわきり じゅんいち)

- 2006年 宮崎大学工学部情報システム工学科卒業
- 2008年 宮崎大学大学院工学研究科修士課程修了 (情報システム工学専攻)
- 2012年 宮崎大学大学院医学系研究科博士後期課程修了 (医学専攻), 博士 (医学)
- 2012年 東京大学特任研究員
- 2018年 東京大学特任助教
- 2019年 東京大学助教
- 現在に至る

# バイオインフォマティクスのための生命科学入門

An Introduction to Life Sciences for Bioinformatics

© Tsukasa Fukunaga, Junichi Iwakiri 2022

2022年8月3日 初版第1刷発行

検印省略

監修者 浜田道昭  
著者 福永津嵩  
岩切淳一  
発行者 株式会社 コロナ社  
代表者 牛来真也  
印刷所 三美印刷株式会社  
製本所 株式会社 グリーン

112-0011 東京都文京区千石 4-46-10

発行所 株式会社 コロナ社  
CORONA PUBLISHING CO., LTD.

Tokyo Japan

振替 00140-8-14844 · 電話 (03) 3941-3131(代)

ホームページ <https://www.coronasha.co.jp>

ISBN 978-4-339-02731-0 C3355 Printed in Japan

(新宅)



**JCCOPY** <出版者著作権管理機構 委託出版物>

本書の無断複製は著作権法上での例外を除き禁じられています。複製される場合は、そのつと事前に、出版者著作権管理機構（電話 03-5244-5088, FAX 03-5244-5089, e-mail: info@jcopy.or.jp）の許諾を得てください。

本書のコピー、スキャン、デジタル化等の無断複製・転載は著作権法上での例外を除き禁じられています。購入者以外の第三者による本書の電子データ化及び電子書籍化は、いかなる場合も認めていません。落丁・乱丁はお取替えいたします。