

脊椎動物のゲノムに存在する
ボルナウイルスの分子化石を用いた
古ウイルス学研究

川崎 純菜

目次

| | |
|--|----|
| 要旨 | 3 |
| 略号表 | 5 |
| 第一章 序論 | |
| 1-1. 古ウイルス学と内在性ウイルス様エレメント (EVEs) | 8 |
| 1-2. 内在性ボルナウイルス様エレメント (EBLs) | 12 |
| 1-3. ボルナウイルスの基本性状 | 14 |
| 1-4. 先行研究の問題点と本研究の目的 | 17 |
| 第二章 結果 | |
| 2-1. 真核生物ゲノムにおける EBLs の網羅的探索 | 20 |
| 2-2. 約 1 億年にわたるボルナウイルスの感染の歴史 | 25 |
| 2-3. ボルナウイルスは多様な脊椎動物を宿主としてきた | 31 |
| 2-4. 古代ボルナウイルスの地理的分布 | 32 |
| 2-5. 霊長類祖先に感染していたボルナウイルス系統の時系列変化 | 35 |
| 2-6. 霊長類祖先とボルナウイルスとの数千万年にわたる共存 | 36 |
| 第三章 考察 | |
| 3-1. EVEs を用いた古ウイルス学研究 | 41 |
| 3-2. EBLs を用いた研究の限界 | 42 |
| 3-3. ボルナウイルスの長期的進化の背後にあるメカニズム | 42 |
| 3-4. ウイルスと宿主との長期的共存が感染の拡大をもたらした可能性 | 46 |
| 3-5. 宿主生物ごとのボルナウイルスの内在化の頻度の違い | 48 |
| 3-6. EVEs の活用によるウイルスの多様性の解明 | 49 |
| 第四章 材料と方法 | |
| 4-1. 真核生物のゲノムデータにおける EBLs の探索 | 51 |
| 4-2. EBLs の挿入年代の推定 | 53 |
| 4-3. ネットワーク解析による年代推定法の検証 | 58 |

| | |
|-------------------------------------|----|
| 4-4. 系統解析..... | 61 |
| 4-5. EBLs の由来となった古代ボルナウウイルスの分類..... | 63 |
| 4-6. 古代ボルナウウイルスの遺伝的多様性の評価..... | 63 |
| 4-7. データの公開 | 64 |
| 引用文献..... | 65 |
| 注釈..... | 72 |
| 謝辞..... | 73 |

要旨

生物は進化の過程において、ウイルスに繰り返し感染してきたと考えられる。しかし、ウイルスは物理的化石を残すことがないため、先史時代におけるウイルス感染を研究することは困難であると考えられてきた。生物のゲノムには、過去に感染したウイルスに由来する遺伝子配列である内在性ウイルス様エレメント (Endogenous Viral Elements: EVEs) が存在する。古ウイルス学は、生物のゲノムに存在する EVEs を「ウイルスの分子化石」として解析し、過去のウイルスの宿主生物や存在年代を明らかにすることで、ウイルスの感染の歴史を追究する学問である。

内在性ボルナウイルス様エレメント (Endogenous Bornavirus-Like elements: EBLs) は、RNA ウイルスの一種であるボルナウイルスに由来する EVEs であり、さまざまな生物ゲノムにおいて同定されている。したがって、EBLs を用いた研究によりボルナウイルス感染の歴史を詳細に追跡することができる。本研究では、「ボルナウイルスの分子化石」である EBLs を 969 種の真核生物のゲノムにおいて網羅的に探索し、遺伝子オーソロジー解析により古代ボルナウイルスの宿主域や感染年代を推定することで、ボルナウイルス感染の歴史の解明を試みた。

EBLs の網羅的解析により、1)約 1 億年前の白亜紀にはボルナウイルスの感染がすでに発生していたこと、2)その後もボルナウイルスの感染が多様な脊椎動物において発生していたことを明らかにした。これらの結果は、白亜紀から現代に至るまで、ボルナウイルスが流行を繰り返しながらも脊椎動物と共存してきたことを示している。

さらに本研究では、霊長類の進化過程において、ボルナウイルスの内在化が何度も繰り返し発生してきたことを見出した。EBLs を用いた系統解析により、1)霊長類の祖先には少なくとも 6 つの異なる系統のボルナウイルスが感染してきたこと、2)これらのボルナウイルス系統は異なる時代に霊長類祖先のゲノムに挿入されたことを明らかにし、霊長類の祖先におけるボルナウイルスの流行が時代とともに変化してきたことを示唆した。

一方で、系統解析により、数千万年にわたって霊長類祖先に内在化を繰り返していたボルナウイルス系統を同定し、これらのボルナウイルスが霊長類祖先と長期的な共存関係を築いていた可能性を示した。また、霊長類祖先と長期的に

共存していたボルナウイルスに由来する EBLs が、複数種のコウモリのゲノムにおいても検出されたことから、この系統のボルナウイルスが霊長類の祖先だけでなく、コウモリの祖先にも感染していたことが示された。さらに、祖先生物の生息域についてのこれまでの知見と本研究で得られたボルナウイルスの内在化時期を統合的に解析した結果、1)霊長類祖先は古代ボルナウイルスと共存していた時代に複数の大陸間を移動していたこと、2)この時代におけるコウモリ祖先の生息域は霊長類祖先の移住経路と大陸レベルで重なっていたことを明らかにした。以上より、長期的なウイルスと宿主生物の共存関係が、ボルナウイルスの地理的分布の拡大、さらにはコウモリへのウイルス伝播に関与した可能性を示した。本研究は、脊椎動物の進化過程におけるボルナウイルス感染の歴史を再構築し、生物進化における RNA ウイルスの流行と共存の歴史の理解に貢献した。

略号表

本文中および図表中で用いた略号は、以下のとおりである。

| | |
|--------|--|
| BLAST | : Basic local alignment search tool |
| DNA | : Deoxyribonucleic acid |
| dsDNA | : Double-stranded DNA |
| dsRNA | : Double-stranded RNA |
| EBLG | : Endogenous bornavirus-like G element |
| EBLL | : Endogenous bornavirus-like L element |
| EBLM | : Endogenous bornavirus-like M element |
| EBLN | : Endogenous bornavirus-like N element |
| EBLP | : Endogenous bornavirus-like P element |
| EBL | : Endogenous bornavirus-like element |
| EVE | : Endogenous viral element |
| G | : Glycoprotein |
| ICTV | : International Committee on Taxonomy of Viruses |
| INSDB | : International Nucleotide Sequence Databases |
| kb | : Kilo base pairs |
| L | : RNA-dependent RNA polymerase |
| LINE-1 | : Long interspersed nuclear element-1 |
| M | : Matrix protein |
| mRNA | : Messenger RNA |
| MSA | : Multiple sequence alignment |
| MYA | : Million years ago |
| N | : Nucleoprotein |
| NCBI | : National Center for Biotechnology Information |
| ND | : Not detected |
| ORF | : Open reading frame |
| P | : Phosphoprotein |
| PolyA | : Polyadenylic acid |
| RNA | : Ribonucleic acid |
| RT | : Reverse transcription |

| | |
|----------|---------------------------------------|
| ssDNA | : Single-stranded DNA |
| TE | : Transposable element |
| TSD | : Target site duplication |
| UCSC | : University of California Santa Cruz |
| X | : X protein |
| (+)ssRNA | : Positive-sense single-stranded RNA |
| (-)ssRNA | : Negative-sense single-stranded RNA |
| % | : Percent |

第一章

序論

1-1. 古ウイルス学と内在性ウイルス様エレメント (EVEs)

古ウイルス学は過去のウイルスの感染を研究する分野である。生物は進化の過程において、ウイルスの脅威に繰り返し晒されてきたと考えられる。しかし、ウイルスは物理的な化石を残すことなく消失してしまうため、過去のウイルスの感染を調査することは困難であると考えられてきた。近年、生物の化石や博物館に保管されていた標本からのウイルス遺伝子配列の検出により、過去のウイルスの感染調査が可能になりつつある(1-6)。ただし、こうした調査では最大でも数百年前から数千年前に発生したウイルスの感染までしか遡ることができず、それ以前のウイルス感染の歴史については謎に包まれたままとなっていた。

生物のゲノムには、過去に感染したウイルスの遺伝子配列に由来する、内在性ウイルス様エレメント (EVEs) が存在している。EVEs は、ウイルスの遺伝子配列が宿主生物の生殖細胞ゲノムに挿入され、生物のゲノムの一部として子孫へと受け継がれてきた結果だと考えられている(7)。つまり、EVEs を「ウイルスの分子化石」として解析することで、過去のウイルスの存在年代や宿主生物を明らかにし、数千万年から数億年にわたるウイルスの感染の歴史を調査することができる(7-11)。

ウイルスはゲノム核酸の性状と複製の様式により、DNA ウイルス、RNA ウイルス、逆転写 RNA ウイルス、逆転写 DNA ウイルスの 4 つのグループに分類される (図 1A)。このうち、RNA ウイルスは RNA のみを介して複製を行うのに対して、他の 3 つのグループに属するウイルスは DNA を中間体とした複製を行う(12)。

これまでの研究により、前述した 4 つのグループのウイルスに由来する EVEs が生物のゲノムにおいて同定されている (図 1B) (7, 8, 13, 14)。最初に発見された EVEs は、逆転写 RNA ウイルス (レトロウイルス) に由来するものである(15)。レトロウイルスは、自身がコードする逆転写酵素とインテグラーゼにより、ウイルスの一本鎖 RNA ゲノムを二本鎖 DNA へと逆転写し、宿主生物ゲノムへと挿入する。つまり、レトロウイルスは、自律的にウイルス配列を宿主生物のゲノムへと挿入することができる。そのため、レトロウイルスに由来する EVEs は生物のゲノムの高い割合を占めており、ヒトゲノムの約 8%、マウスゲノムの約 10%を構成している(16, 17)。

対して、レトロウイルス以外のウイルスは、自身の遺伝子配列を宿主生物のゲノムに挿入する能力を持たない。それにもかかわらず、こうしたウイルスに由来する遺伝子配列は生物のゲノムにおいて多数同定されている(7, 8, 13, 14)。DNA ウイルスおよび逆転写 DNA ウイルス（ヘパドナウイルス）の場合は、非相同末端結合や相同組換えにより、ウイルスの DNA 配列が宿主生物ゲノムに挿入されると考えられている(18-23)。RNA ウイルスの場合は、レトロトランスポゾンがコードする逆転写酵素により、ウイルスの RNA 配列が DNA 配列へと変換され、宿主生物のゲノムに挿入されると考えられている(14, 24)。

レトロトランスポゾンは、生物のゲノムに存在する転移因子（TEs）の一種であり、自身の DNA 配列から転写された RNA 配列を、逆転写反応により DNA にコピーし、このコピー配列を新たなゲノム位置へと挿入する(25)。RNA ウイルスの内在化には、いくつかのタイプのレトロトランスポゾンが関与していると考えられている(26)。なかでも、先行研究における配列解析および実験検証の結果から、long interspersed nuclear element-1（LINE-1）の関与が強く示唆されている(7, 8, 14, 27-29)。LINE-1 は RNA 結合タンパク質である ORF1 タンパク質と、エンドヌクレアーゼ活性および逆転写酵素活性をもつ ORF2 タンパク質をコードする(25)。LINE-1 は自身の RNA だけでなく、他のレトロトランスポゾンや細胞由来の mRNA の逆転写および生物ゲノムへの挿入にも関与する(30-32)。LINE-1 による逆転写と挿入の効率には、RNA の 3'末端に存在する polyA 配列が大きく影響することが報告されている。また、LINE-1 による挿入配列の両側には、標的部重複（Target Site Duplication: TSD）と呼ばれる短いリピート配列が形成される。以上より、LINE-1 による挿入配列の周囲には polyA 配列および TSD といった特徴が存在することが知られている。こうした特徴は、RNA ウイルスに由来する EVEs の挿入部位においても観察されている(7, 8, 14, 27-29)。加えて、RNA ウイルスの一種であるボルナウイルスを用いた感染実験の結果、ウイルスの配列が生物のゲノムに挿入され、さらに挿入されたウイルス配列の周辺には polyA 配列および TSD といった LINE-1 による挿入を示唆する特徴が存在していたことが報告されている(14, 27)。このような結果は、LINE-1 が RNA ウイルスの内在化に関与している可能性を示している。一方で、RNA ウイルスの内在化に関わるプロセスの詳細については、未解明な点が多く残されている（詳細は「第三章 考察」に記述する）。

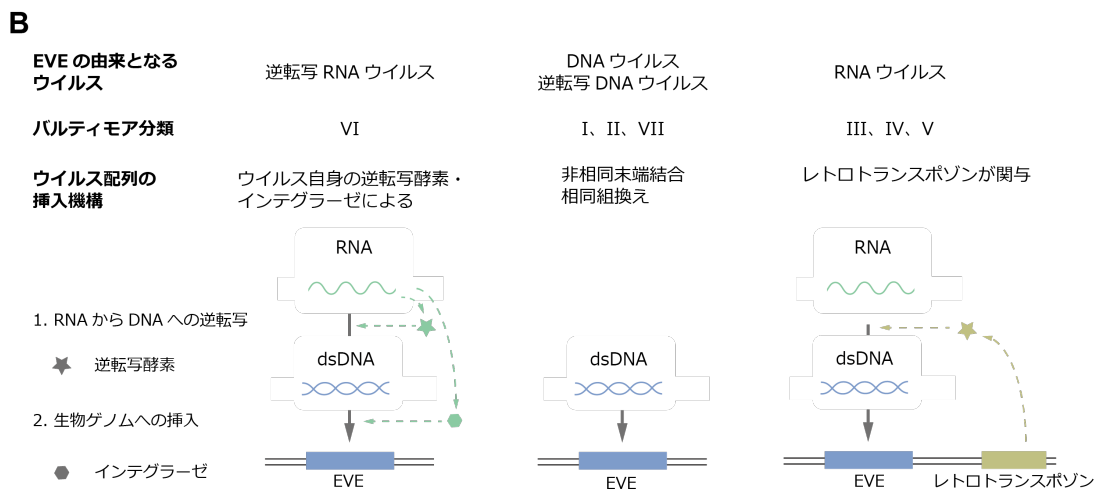
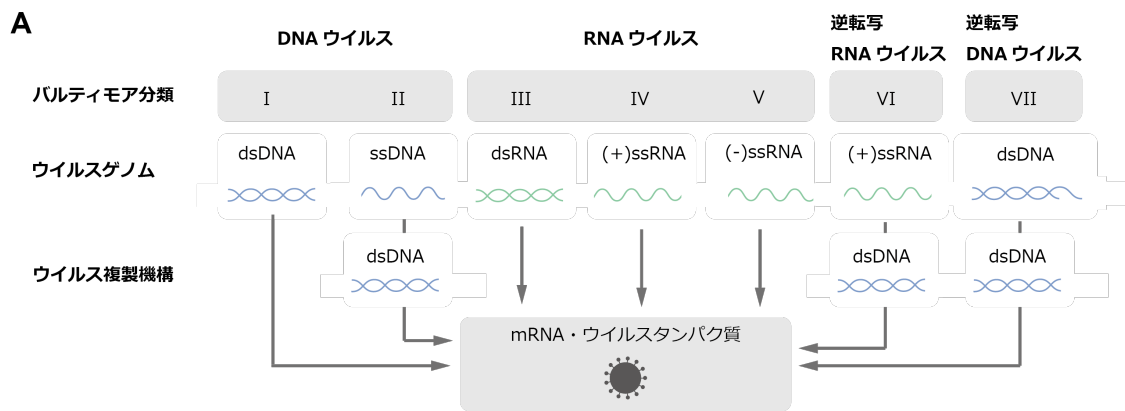


図 1. ウイルスの複製、およびウイルス配列の生物ゲノムへの挿入

(A) ウイルスの複製機構の模式図。ウイルスは、ゲノム核酸の性状と複製における逆転写反応の有無により、DNA ウイルス、RNA ウイルス、逆転写 RNA ウイルス、逆転写 DNA ウイルスの 4 つのグループに分類される。また、バルティモア分類により、ウイルスは以下のように細分化される (I: 二本鎖 DNA ウイルス、II: 一本鎖 DNA ウイルス、III: 二本鎖 RNA ウイルス、IV: 一本鎖プラス鎖 RNA ウイルス、V: 一本鎖マイナス鎖 RNA ウイルス、VI: 一本鎖 RNA 逆転写ウイルス、VII: 二本鎖 DNA 逆転写ウイルス)。逆転写 DNA ウイルス (ヘパドナウイルス) は、部分的に一本鎖構造をもつ不完全な二本鎖 DNA をゲノムとして持ち、ウイルスゲノムを複製する過程で逆転写反応を行う。(B) ウイルス配列の宿主生物ゲノムへの挿入機構の模式図。逆転写 RNA ウィル

ス（レトロウイルス）は、逆転写酵素とインテグラーゼをコードし、自律的にウイルス配列を生物のゲノムに挿入する。DNA ウイルスおよび逆転写 DNA ウイルス（ヘパドナウイルス）の場合は、非相同末端結合や相同組換えを介して、ウイルス配列が生物ゲノムへと挿入されると考えられている。RNA ウイルスの場合は、レトロトランスポゾンがコードする逆転写酵素により、ウイルスに由来する RNA 配列が DNA 配列へと変換され、その後、生物ゲノムに挿入されると考えられている。図中では、レトロトランスポゾンによる逆転写と挿入が段階的に進行する過程が示してあるが、レトロトランスポゾンの種類によっては、逆転写と挿入がほぼ同時に起こる。

1-2. 内在性ボルナウイルス様エレメント (EBLs)

現在、多様な生物のゲノムが解読され、さまざまなウイルスに由来する EVEs が同定・報告されている。しかし、ウイルス学の歴史において、RNA ウイルスに由来する EVEs の存在は、ほとんど研究されてこなかった。この理由は、前項「1-1. 古ウイルス学と内在性ウイルス様エレメント (EVEs)」で記載したとおり、RNA ウイルスは複製の過程で DNA を形成せず、通常の RNA ウイルスの感染において、ウイルスの遺伝子配列が生物のゲノムに挿入されることは起こり得ないと考えられていたためである。一方で、ウイルス感染実験において RNA ウイルスの遺伝子配列が DNA に逆転写されることや、逆転写された DNA が生物のゲノムに挿入されることが報告されていた(24, 33, 34)。ただし、こうした現象は培養細胞または実験動物を用いた短期的な実験において観察されたものであり、RNA ウイルス配列の挿入が体細胞においてのみ起こりうるのか、または生殖細胞においても発生し、さらには RNA ウイルス配列が子孫生物へと受け継がれうるのかといった点は不明なままであった。

こうしたなか、RNA ウイルスの一種であるボルナウイルスに由来する EVEs (内在性ボルナウイルス様エレメント: EBLs) が、ヒトを含むさまざまな哺乳類のゲノムに存在することが明らかとなった(ボルナウイルスについては、次項「1-3. ボルナウイルスの基本性状」で詳述する)(14)。その後、生物のゲノムにおける EVEs の網羅的探索に関する研究が相次いで報告され、ボルナウイルスだけでなく、さまざまなウイルスに由来する EVEs が広く生物のゲノムに存在していることが示された(図 2)(7, 8, 14, 27-29)。こうした成果により、過去に感染した RNA ウイルスの遺伝子配列が生物のゲノムに内在化し、現代に至るまで子孫生物へと受け継がれていることが示された。

興味深いことに、前述した EVEs の網羅的探索により、脊椎動物のゲノムにおいて最も頻繁に検出される「RNA ウイルスの分子化石」は EBLs であることが示されている(図 2)。つまり「ボルナウイルスの分子化石」は、脊椎動物のゲノムに存在する「RNA ウイルスの分子化石」のなかで、最もデータ量が多いものにあたる。したがって、EBLs を用いた研究により、脊椎動物の進化過程におけるボルナウイルスの感染を詳細に調査することができるだけでなく、RNA ウイルスの感染の歴史について、1つのモデルを提供することができると期待される。

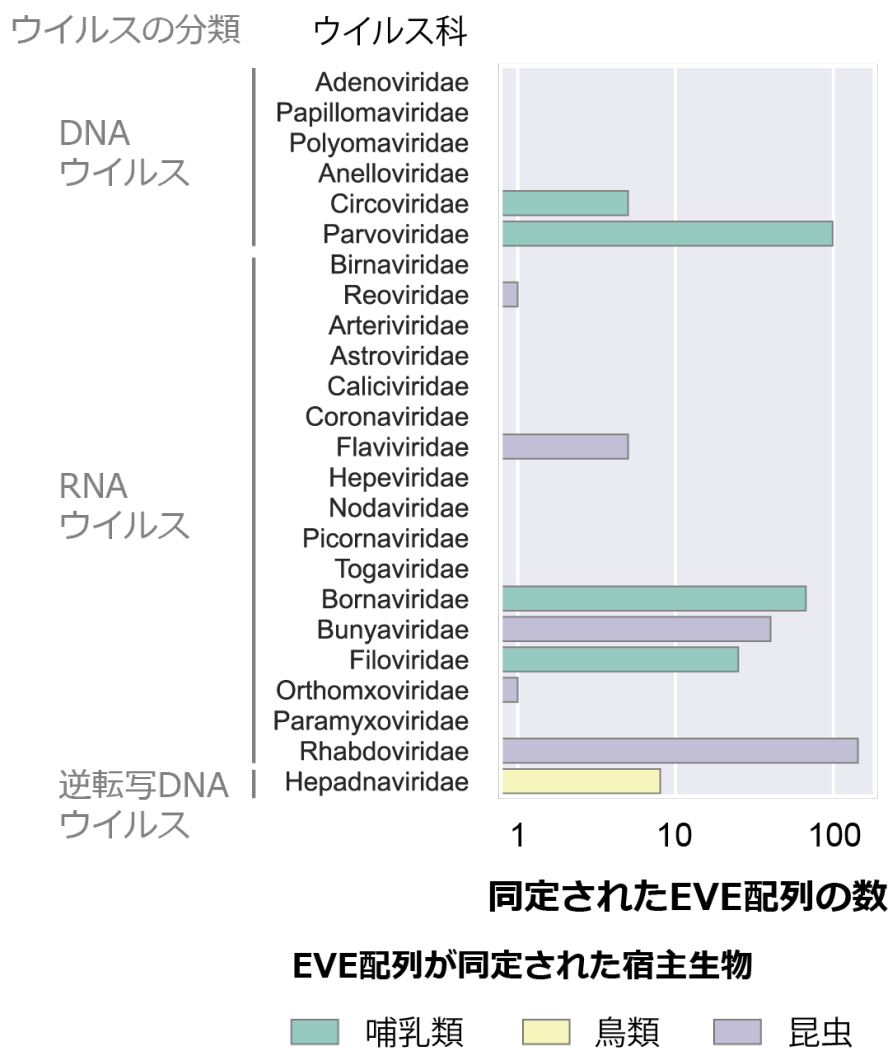


図 2. 生物のゲノムにおける EVEs の網羅的探索

本図は、Katzourakis and Gifford., *Plos Genetics*, 2010 (8)において発表されたデータを使用して作成した。Y 軸は EVEs の由来となったウイルスを、X 軸は同定された EVE 配列の数を示している。棒グラフの色は EVE 配列が同定された宿主生物を示す：哺乳類（緑）、鳥類（黄）、昆虫（紫）。

1-3. ボルナウイルスの基本性状

ボルナウイルスは、一本鎖マイナス鎖 RNA をゲノムにもつモノネガウイルス目に属するウイルスである。現在、ボルナウイルス科にはオルソボルナウイルス属、カルボウイルス属、カルターウイルス属が存在する(35)。2018年まで、ボルナウイルス科は単一のウイルス属(今日のオルソボルナウイルス属、旧名はボルナウイルス属)により構成されていた。しかし、新種のボルナウイルスがニシキヘビ、およびコイ科の魚において発見されたことにより、カルボウイルス属、カルターウイルス属が設立された(36, 37)。

ボルナウイルスのゲノムには、少なくとも6つの open reading frames(ORFs)が存在し、ヌクレオプロテイン(N)、Xタンパク質(X)、リン酸化タンパク質(P)、マトリックスタンパク質(M)、糖タンパク質(G)、RNA依存性RNAポリメラーゼ(L)がコードされている(図3)。オルソボルナウイルスのゲノムでは、これらのウイルス遺伝子が3'末端よりN、X、P、M、G、Lの順に配置されているが、カルボウイルス、カルターウイルスのゲノムでは、N、X、P、G、M、Lの順に配置されている。

それぞれのウイルスタンパク質の機能は、オルソボルナウイルスを用いた研究によって明らかにされてきた(38, 39)。N、P、Lはポリメラーゼ複合体を構成し、ウイルスゲノムの複製とmRNAの転写を制御する。このうち、Nはウイルスゲノムに結合するRNA結合タンパク質であり、PはポリメラーゼであるLの補因子として機能する。Xはウイルスのポリメラーゼ活性を抑制する機能が報告されている。MとGは構造タンパク質である。Gはウイルスの宿主細胞への吸着や侵入に関与するエンベロープタンパク質であり、Mはウイルス粒子を裏打ちするタンパク質であると考えられている。これまでの研究により、N、M、G、L遺伝子に由来するEBLsが同定されており、それぞれ、EBLN、EBLM、EBLG、EBLLと名付けられている(7, 8, 14, 27-29)。

ボルナウイルスは多様な脊椎動物に感染することが報告されている(35)。オルソボルナウイルス属に分類されるウイルスは、哺乳類や鳥類、爬虫類を宿主とする。オルソボルナウイルスの一種であるボルナ病ウイルスは、神経系組織に好んで感染し、ウマやヒツジに脳脊髄炎を引き起こすことが知られている。また近年、ヒトに致死性の脳炎を引き起こす新種のウイルスの発見や(40-42)、脳炎患者におけるボルナ病ウイルスの感染が報告されたことにより(43-46)、オルソ

ボルナウイルスは新たな人獣共通感染症の原因として注視されている。また、オウム目やスズメ目の鳥類に感染するオルソボルナウイルスは、腺胃拡張症に代表される消耗性の神経疾患を引き起こす(47)。

対して、カルボウイルスやカルターウイルスの感染性や病原性はほとんど明らかになっていない(35)。カルボウイルスの感染は、弛緩性麻痺や軽度から中等度の脳炎症状を示すニシキヘビ (*Morelia spilota*) から同定されたが、ウイルスの感染とこうした臨床症状との因果関係は不明のままとなっている(36)。またカルターウイルスの感染は、中国およびオーストラリアで採取されたコイ科の魚 (*Hemiculter leucisculus* および *Cyprinus carpio*) において同定されているが、どちらの宿主生物も顕著な病状を示していなかったと報告されている(37, 48)。

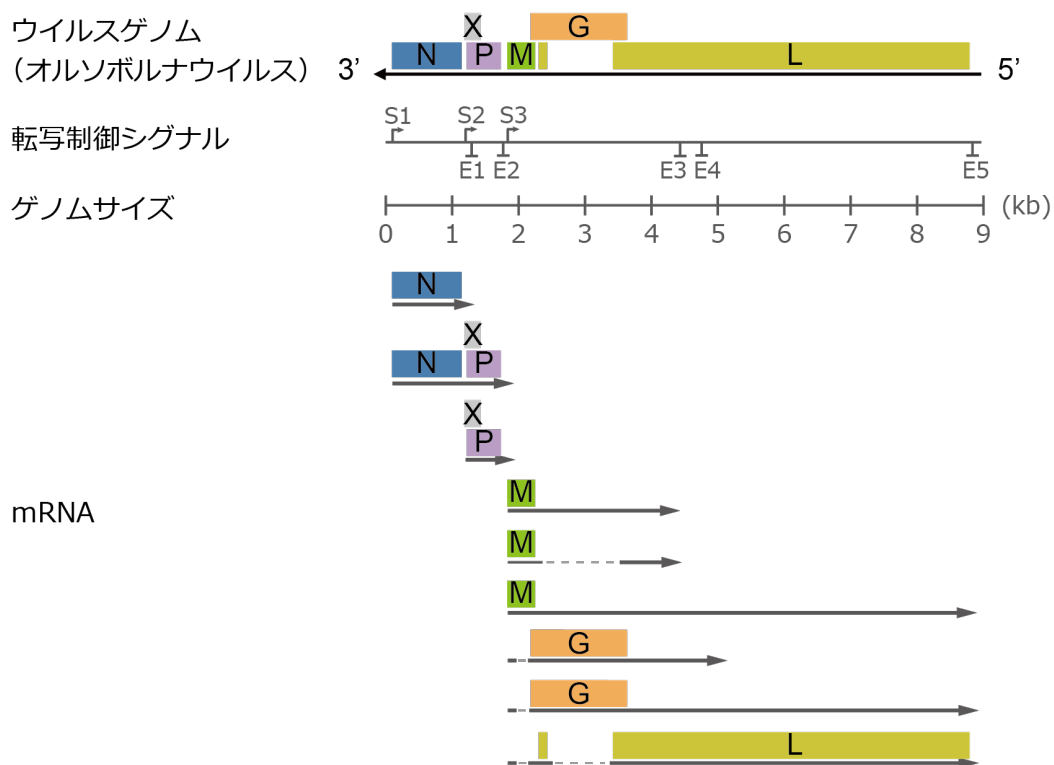


図 3. ボルナウイルスのゲノム構造と転写産物

本図では、オルソボルナウイルスのゲノム構造と転写産物を示す。ボルナウイルスのゲノムには、以下の6つウイルスタンパク質をコードする ORFs が存在する：ヌクレオプロテイン (N)、X タンパク質 (X)、リン酸化タンパク質 (P)、マトリックスタンパク質 (M)、糖タンパク質 (G)、RNA 依存性 RNA ポリメラーゼ (L)。それぞれのウイルスタンパク質は、転写開始シグナル (S1 から S3) および転写終止シグナル (E1 から E5) によって制御される mRNA から生成される。またオルソボルナウイルスは、宿主生物の RNA スプライシング機構により、複数の転写産物アイソフォームを発現することが報告されている(49-52)。図中の mRNA の灰色の点線部分はイントロン領域を示す。EBLs は、ボルナウイルスの mRNA がレトロトランスポゾンにより逆転写され、宿主生物のゲノムへと挿入された結果であると考えられており、これまでに、N、M、G、L 遺伝子に由来する EBLs が同定されている。

1-4. 先行研究の問題点と本研究の目的

EBLs の同定は、宿主生物のゲノムにおいて、現代のボルナウイルスの遺伝子配列に類似した領域を探索することにより行われる(7, 26)。これまでに、さまざまな生物のゲノムにおいて EBLs の存在が報告されているが、こうした EBLs はオルソボルナウイルスの配列のみを用いた探索により同定されてきた(図 4)(8, 14, 28)。つまり、先行研究ではカルボウイルスやカルターウイルスに類似した EBLs は検出できておらず、ボルナウイルスの感染の歴史を部分的にしか捉えられていなかった可能性が高い。実際に、カルボウイルスの遺伝子配列を用いた探索により、新たな EBLs が同定されている(36)。

カルボウイルスとカルターウイルスの発見により、これまで見過ごされてきた EBLs の存在が示唆された。しかし、これらのウイルスに由来する EBLs を網羅的に同定・解析した研究はなく、過去のボルナウイルスの感染についての理解は不十分なままとなっている。そこで本研究では、ボルナウイルスの感染の歴史を再構築することを目的とし、3つのボルナウイルス属に由来する EBLs を網羅的に同定するための大規模解析を実施した。はじめに、969種の真核生物のゲノムデータを用いて EBLs の探索を行った。次に、EBLs の宿主生物ゲノムへの挿入年代を推定することで、EBLs の由来となった古代ボルナウイルスの存在年代と宿主生物を明らかにし、ボルナウイルスと宿主生物との共進化過程の描出を試みた。

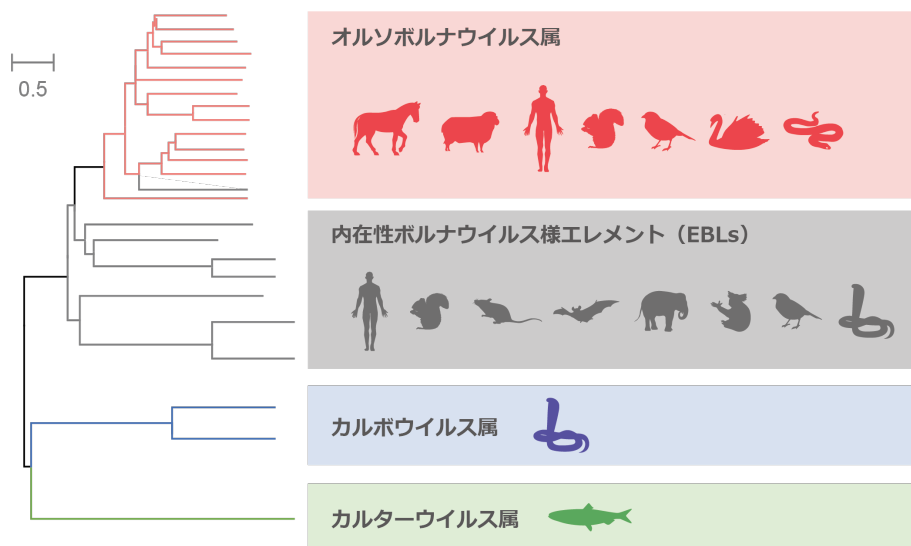


図 4. ボルナウイルス科に分類されるウイルスの系統関係

本図は、国際ウイルス分類委員会 (ICTV) による提案書(53)のデータを改変したものである。(左図) ボルナウイルス科に分類されるウイルスの系統樹。枝の色は以下のとおりである: オルソボルナウイルス (赤)、カルボウイルス (青)、カルターウイルス (緑)、EBLs (灰)。スケールバーは遺伝的距離 (座位あたりの置換数) を示す。先行研究で同定された EBLs は、オルソボルナウイルスに類似したものばかりであった。(右図) それぞれのボルナウイルスの代表的な宿主生物。EBLs の場合は、これまでに EBLs が同定された生物種の代表例を示している。

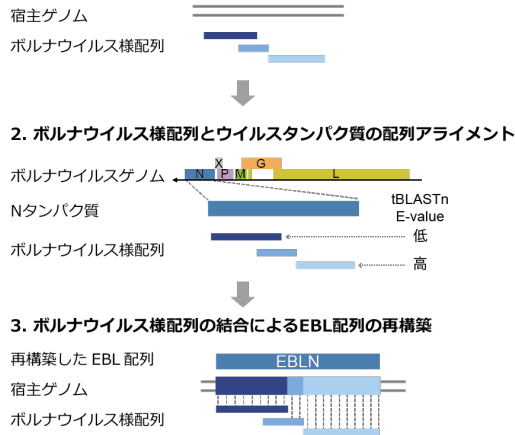
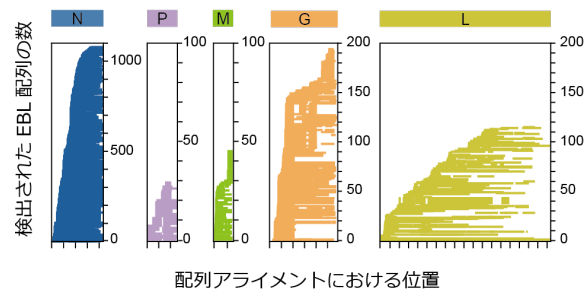
第二章

結果

2-1. 真核生物ゲノムにおける EBLs の網羅的探索

EBLs を網羅的に同定するために、969 種の真核生物のゲノムデータにおいて、tBLASTn によりボルナウイルス様配列を探索した (図 5A)。探索には、ボルナウイルス科の 3 つのウイルス属の配列をクエリとして使用した。次に、ボルナウイルス様配列の宿主ゲノムにおける位置、および現代のボルナウイルスタンパク質とのアライメント位置に基づいて、EBL 配列を再構築した。これは、EBLs の多くが、フレームシフト変異により断片化された配列として検出されていたためである。

EBLs の探索と配列の再構築により、1,465 の EBL 配列を 131 種の脊椎動物において同定した (図 5B)。先行研究により、EBLN、EBLM、EBLG、EBLL の存在は報告されていたが(8, 14, 26)、今回の探索により、新たに 30 の EBLP 配列を発見した。

A**1. 宿主生物のゲノムにおけるボルナウイルス様配列の探索****B****図 5. 真核生物ゲノムにおける EBLs の探索**

(A) EBLs の探索および配列の再構築の方法。まず、真核生物のゲノムにおいてボルナウイルス様配列を tBLASTn により探索した。次に、ボルナウイルス様配列と現代のボルナウイルスのタンパク質との配列アライメントを作成した。最後に、1)ボルナウイルス様配列が検出された宿主ゲノム位置と、2)ウイルスタンパク質とのアライメント位置に基づいてボルナウイルス様配列を結合し、EBL 配列を再構築した。同じゲノム位置で複数のボルナウイルス様配列が検出された場合には、tBLASTn において低い E-value で検出された配列を信頼性の高い配列と見なし、EBL 配列の再構築に使用した。詳細は「第四章 材料と方法」を参照。(B)現代のボルナウイルスのタンパク質と EBLs との配列アライメントの模式図。それぞれのパネルは、EBLs の由来となったウイルス遺伝子に対応する。X 軸は、配列アライメントにおける位置を表しており、100 アミノ酸ごとに目盛線が打たれている。Y 軸は本研究において検出された EBL 配列の数を示している。

次に、EBLs の由来となった古代ボルナウイルスをウイルス属レベルに分類するために、系統解析を行った。系統解析に使用した EBLs はすべて、現代のウイルス属に対応する 3 つのクレードに分割された (図 6)。この結果から、今回同定した EBLs のなかに未知のウイルス属に由来する配列が含まれる可能性は低いと考えられた。同定した 1,465 EBLs のうち、692 配列は系統解析によりウイルス属レベルに分類することができたが、その他の EBLs は配列長が短く、系統解析に使用することができなかった。そこで、EBLs と現代のボルナウイルスとの配列類似性を活用することで、すべての EBLs をウイルス属レベルに分類しようと試みた。まず、系統解析による分類との比較により、配列類似性による分類の正確性を検証した。その結果、2 つの分類法の一致度は EBLNs では 99.7%、EBLGs では 100%、EBLLs では 100%であった。配列類似性に基づく分類が系統解析による分類と高度に一致していることが確認されたため、次に、本研究で同定したすべての EBLs を現代ボルナウイルスとの配列類似性に基づき 3 つのウイルス属に分類した。その結果、同定した EBLs のうち、364 配列はオルソボルナウイルス属、729 配列はカルボウイルス属、372 配列はカルターウイルス属に由来すると推定された (図 7)。

注目すべきは、同定した 1,465 の EBLs のうち、870 もの配列がカルボウイルスおよびカルターウイルスを用いた探索によって初めて検出された点である。これは、先行研究で主に行われてきたオルソボルナウイルス配列のみを用いた探索において、多くの EBLs が見逃されてきたという本研究の仮説を支持する結果である。

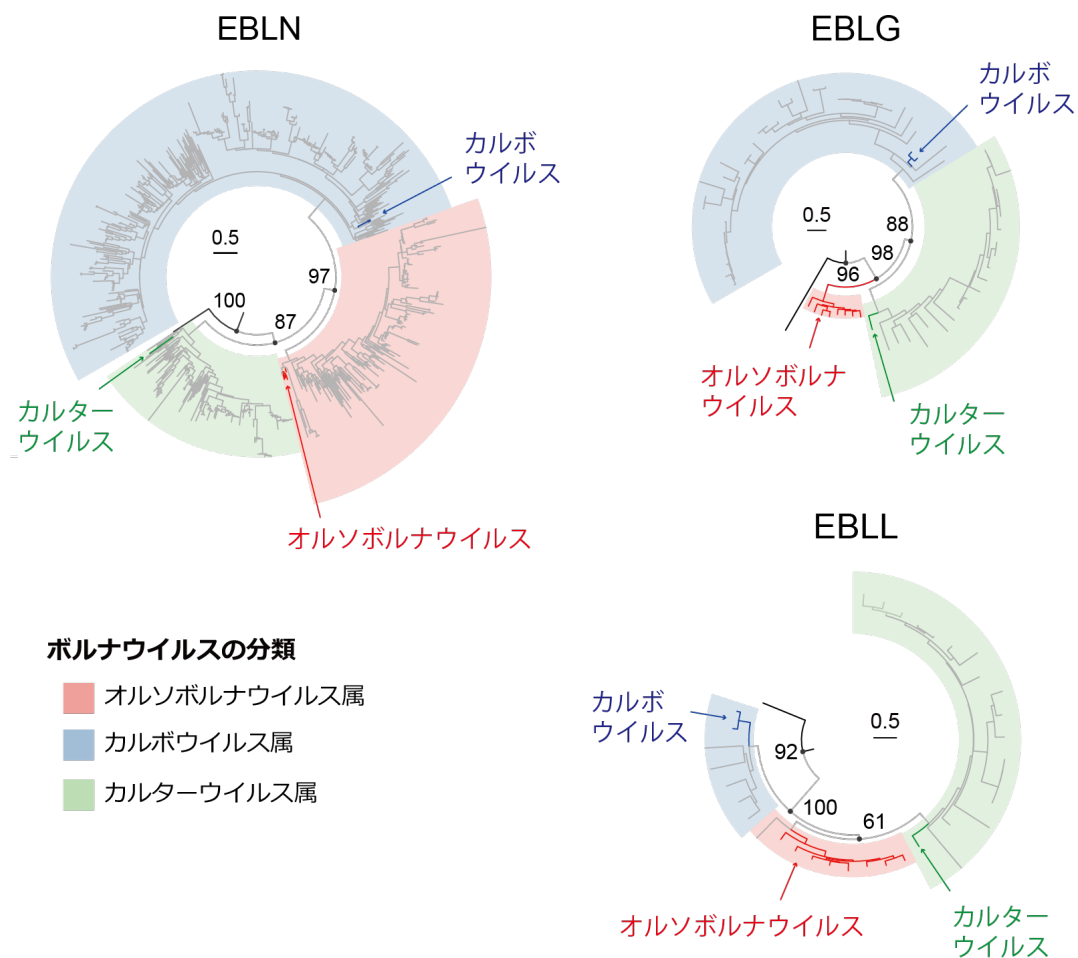


図 6. EBLs と現代のボルナウイルスとの系統解析

これらの系統樹は、EBLs と現代のウイルスタンパク質を用いて、最尤法により構築された。枝の色は以下のとおり対応している：EBLs（灰）、現代のオルソボルナウイルス（赤）、現代のカルボウイルス（青）、現代のカルターウイルス（緑）、現代のニャミウイルス（黒）。これらの系統樹の根の位置は、現代のニャミウイルスを外群とすることによって決定した。ハイライトの色は、現代のボルナウイルスの分類に対応している。枝の支持率(%)はブートストラップ法により算出した。図中では、ウイルス属レベルの分類に関わる枝の支持率のみを示している。スケールバーは遺伝的距離（座位あたりの置換数）を示す。

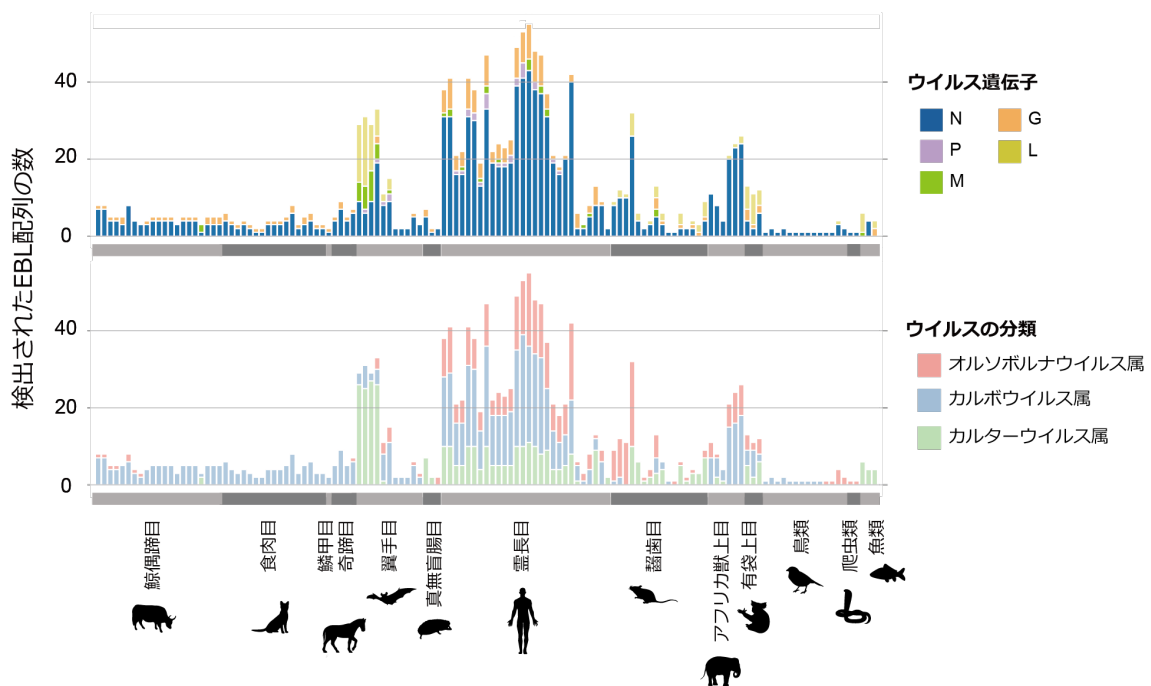


図 7. 宿主生物種ごとの EBLs の検出数

X 軸は EBLs が検出された脊椎動物の種を、Y 軸はその生物種において検出された EBL 配列の数を示す。それぞれのパネルの積み上げ棒グラフは、EBLs の由来となったウィルス遺伝子（上図）、またはウィルスの分類（下図）に基づき色分けされている。

2-2. 約 1 億年にわたるボルナウイルスの感染の歴史

本研究では、EBLs の由来となった古代ボルナウイルスの遺伝子配列がいつごろ宿主生物のゲノムに挿入されたかを推定するために、遺伝子オーソロジー解析を実施した(7)。今回は、大量の EBL 配列の挿入年代を効率的に推定するために、配列類似性ネットワークを用いたオーソログ決定法を開発した(図 8)。この方法では、まず、EBLs の挿入部位間の配列類似性(アライメントカバレッジ)をペアワイズに算出する。次に、配列類似性ネットワークを構築し、コミュニティ構造を抽出することで、EBL 配列をオーソロググループに分割する。最後に、オーソログを保有する宿主生物の種分岐の年代に基づき、EBLs の挿入年代を推定する。

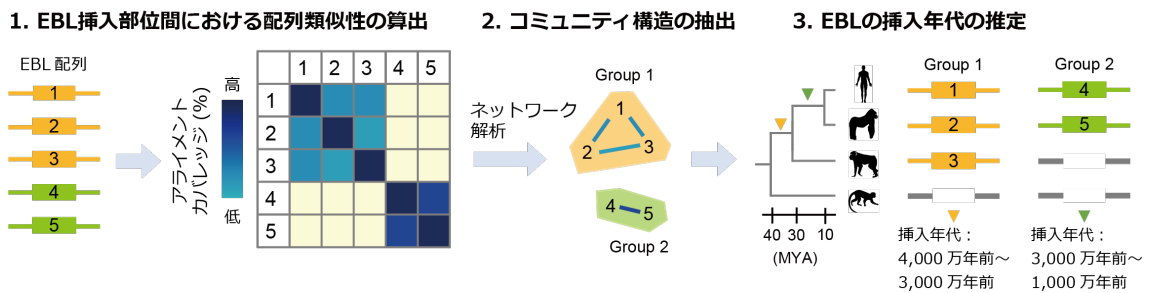
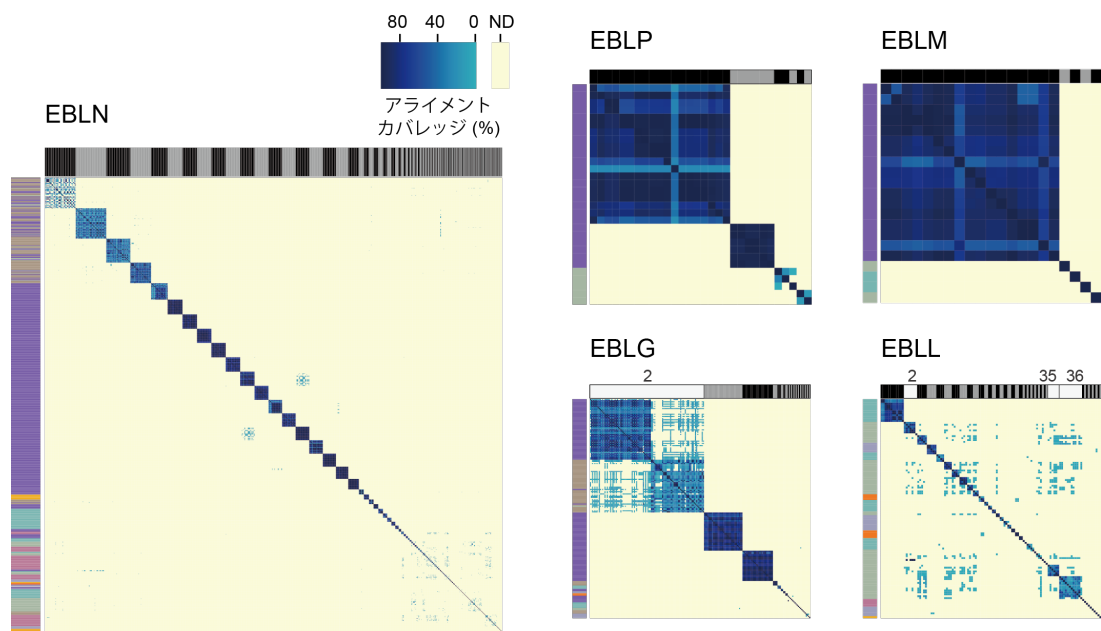


図 8. EBLs の挿入年代の推定方法

ここでは、2つの異なる挿入イベントに由来する5つのEBL配列を想定する。EBL配列の色は、由来となった挿入イベントを示している。はじめに、EBLsの挿入部位の配列類似性をBLASTNによりペアワイズに算出した。ヒートマップの行と列はEBL配列のインデックスに対応しており、セルの色はEBL挿入部位間の配列類似性（アライメントカバレッジ）を示している。アライメントカバレッジは、クエリ配列長あたりのアライメントされたサイトの割合である。次に、EBL配列をオーソロググループに分割するために、1.で作成したヒートマップを入力とした配列類似性ネットワークを構築し、コミュニティ構造の抽出を行った。ネットワークのノードはEBL配列のインデックスに、エッジの色は1.で算出したアライメントカバレッジに対応している。ハイライトは、コミュニティとして抽出されたグループを示しており、本解析ではこれらのコミュニティをオーソロググループと定義した。最後に、EBLsの挿入年代を宿主生物の種分岐の年代に基づき推定した。左側の系統樹は宿主生物の進化過程を示しており、系統樹上の矢頭はオーソログの有無に基づき推定されたそれぞれの挿入イベントの発生年代を示す。系統樹の右側には、オーソログなEBLsの配列アライメントの模式図を示している。MYA (million years ago)：百万年前。詳細は「第四章 材料と方法」を参照。

図 8 に示す遺伝子オーソロジー解析により、1,465 の EBL 配列を 281 のオーソロググループに分割した。図 9 において、ネットワーク解析により分割されたオーソロググループが EBL 挿入部位間の配列類似性を高度に反映していることが確認された。281 のオーソロググループのうち 113 のグループは、複数の宿主生物において検出された EBLs により構成されていたため、生物の種分岐の年代に基づき挿入年代を推測することができた（図 10B「オーソログ検出(+）」）。一方で、168 のグループは単一の宿主生物において検出された EBL 配列のみで構成されていた（図 10B「オーソログ検出(-）」）。単一の宿主生物においてしか検出されなかった EBLs に関しては、2 つの可能性が考えられる。1 つ目には、これらの EBLs は宿主生物が近縁種と分岐した後に挿入された遺伝子座である可能性が考えられる。2 つ目には、近縁種のゲノムが解読されていないために、オーソログが検出できなかった可能性がある。今回はこれら 2 つの可能性を区別することが困難であったため、図 10A ではオーソログが検出された 113 のグループの挿入年代のみを示している。今後、多様な生物のゲノムが解読されることで、EBLs の挿入年代の推定精度が向上すると期待される(54, 55)。

図 10 の結果から、約 1 億年前の白亜紀にはボルナウイルスの感染がすでに発生していたことを明らかにした。先行研究では、ボルナウイルスの感染は少なくとも約 8,300 万年前にアフリカ獣上目の祖先で発生していたことが報告されていた(8, 56)。しかし本研究により、6 つの EBL 遺伝子座のオーソログが北方真獣類に分類される生物のゲノムに存在することが示され、ボルナウイルスの感染が少なくとも約 9,600 万年前に起こっていたことが明らかとなった。つまり、本解析により最古のボルナウイルスの感染の分子化石を同定することに成功した。また、ボルナウイルスの挿入は、北方真獣類の祖先だけでなく、アフリカ獣上目、テティス獣類、有袋上目、霊長目、齧歯目、スズメ目の祖先において、計 18 回発生していたと推定された（図 10A の N1 から N7）。これらの結果は、白亜紀にはすでに、ボルナウイルスがさまざまな脊椎動物に感染を広げていたことを示唆する。



EBL が同定された宿主生物の分類

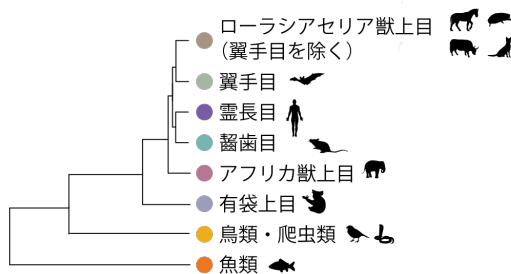


図 9. ネットワーク解析により推定された EBL 配列間のオーソログ関係

ヒートマップの行と列は、本研究で同定した EBL 配列に対応している。ヒートマップのセルの色は EBL 挿入部位間のアライメントカバレッジを示しており、黄色は配列類似性が検出限界以下 (not detected: ND) であったことを意味する。行の色は EBL 配列が同定された宿主生物の分類と対応しており、列の色はネットワーク解析により分類されたオーソロググループを示す。白色の列で示されている EBLG2、EBLL2、EBLL35、EBLL36 は、ネットワーク解析後に再編成されたグループである (詳細は「第四章 材料と方法」を参照)。

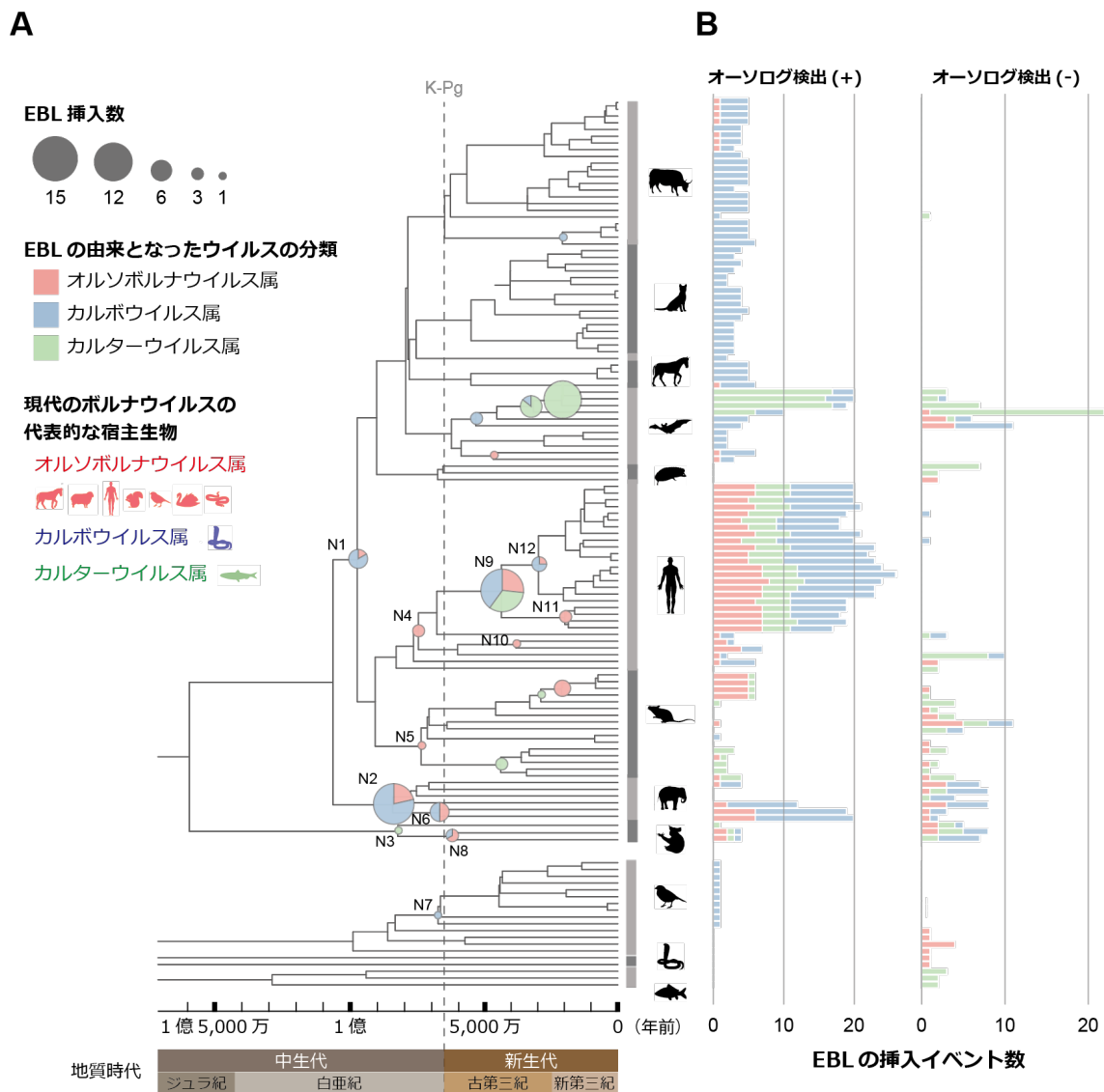


図 10. 約 1 億年にわたる EBLs の挿入の歴史

(A) 脊椎動物の進化過程における EBLs の挿入。系統樹は、脊椎動物の進化過程を示す。系統樹上の円グラフの位置は、いつ、どの生物の祖先において EBLs の挿入が発生したかを示している。円グラフの大きさとは、その時代における EBLs の挿入の数と EBLs の由来となったボルナウイルスの分類を示す。系統樹の内部節における注釈は、以下の宿主生物の共通祖先を示す：北方真獣類 (N1)、アフリカ獣上目 (N2)、有袋上目 (N3)、霊長目 (N4)、齧歯目 (N5)、テティス獣類 (N6)、スズメ目 (N7)、

オーストラリア有袋大目 (N8)、真猿下目 (N9)、キツネザル上科 (N10)、広鼻小目 (N11)、狭鼻小目 (N12)。系統樹上の灰色の点線は、中生代と新生代の境界 (K-Pg 境界) を示す。(B) それぞれの宿主生物の進化の過程で発生した EBLs の挿入イベントの数。Y 軸は (A) の系統樹における宿主生物種と対応しており、X 軸はそれぞれの生物種の進化過程で発生した EBLs の挿入イベントの数を示す。積み上げ棒グラフの色は、(A) と同様、EBLs の由来となったボルナウイルスの分類を表す。ネットワーク解析により分割されたオーソロググループは以下のとおり分類した。「オーソログ検出 (+)」: 複数の宿主生物種において検出された EBLs によりオーソロググループが構成されていた場合、「オーソログ検出 (-)」: 単一の生物種から検出された EBL 配列のみでオーソロググループが構成されていた場合。(A) の系統樹上に示されている円グラフは、「オーソログ検出(+)」のグループの EBLs の挿入年代のみを示している (詳細は「2-2. 約 1 億年にわたるボルナウイルスの感染の歴史」に記述)。

2-3. ボルナウイルスは多様な脊椎動物を宿主としてきた

図 10 の結果から、古代のボルナウイルスは、現代のボルナウイルスが感染することが報告されている宿主生物よりも多様な脊椎動物に感染してきたことが示された。例えば、現代におけるオルソボルナウイルスの感染は、ウマやヒツジ、トガリネズミ、リス、ヒトといった哺乳類(35, 40-46, 57, 58)、オウム目やスズメ目等の鳥類(47)、およびガーターヘビ、ナミヘビ、クサリヘビといった爬虫類(59, 60)で報告されている。対して本研究では、齧歯目、アフリカ獣上目、有袋上目の進化過程において、オルソボルナウイルスの感染が繰り返し発生してきたことを明らかにした(図 10A)。特に、現代のオルソボルナウイルスの疫学調査では、マウスにおけるウイルスの感染は確認されていないにもかかわらず(57, 58)、本研究では、古代のオルソボルナウイルスがマウス祖先に繰り返し感染していたことを示した。同様に、現代のカルボウイルスはニシキヘビ(36)、カルターウイルスはコイ科の魚のみから検出されているが(37, 48)、これらのウイルスの祖先は、哺乳類を含む多様な系統の宿主に感染・内在化してきたことが示された。これらの結果により、ボルナウイルスは、現代のボルナウイルスについての研究から予想されていたよりも、はるかに広範な生物に感染してきたことが示唆された。

2-4. 古代ボルナウウイルスの地理的分布

次に、長期にわたるボルナウウイルスの蔓延状況を描出するために、古代のボルナウウイルスの感染が発生していた地域の推定を試みた。今回は、ボルナウウイルスの挿入が発生した時代における宿主生物の生息地を調査することで、古代ボルナウウイルスの地理的分布を推定した。その結果、古代ボルナウウイルスの感染が複数の大陸で発生していた可能性を示した（図 11）。

まず、中生代に発生したボルナウウイルスの挿入を解析した結果、ローラシア大陸とアフリカ大陸に生息していた祖先生物のゲノムにおいてボルナウウイルスの挿入が発生していたことが示された（図 11A-B）。複数の先行研究により、北方真獣類および霊長目の祖先はローラシア大陸に(61-64)、アフリカ獣上目およびテティス獣類の祖先はアフリカ大陸に生息していたと報告されている(61, 62, 65)。したがって、白亜紀におけるボルナウウイルスの感染は、少なくともこれら2つの大陸で発生していたと考えられた。加えて、オーストラリア有袋大目における EBLs の同定により、ボルナウウイルスの感染が南極大陸またはオーストラリア大陸において発生していた可能性を見出した。現存の有袋類は、南アメリカ大陸およびオーストラリア大陸に生息しているが、これは有袋類の祖先生物が南アメリカ大陸から南極大陸を経由し、オーストラリア大陸へと移住する過程で誕生した種が、それぞれの大陸に定着した結果だと考えられている(66, 67)。本研究では、オーストラリア有袋大目のゲノムには存在するが、南アメリカ大陸に生息する有袋類のゲノムでは検出されない EBLs を同定した（図 11B）。これらの結果から、オーストラリア有袋大目のゲノムに存在する EBLs は、南極大陸またはオーストラリア大陸において発生したボルナウウイルスの感染に由来すると考えられた。

次に、新生代に発生したボルナウウイルスの挿入を解析した結果、霊長類祖先におけるボルナウウイルスの感染が複数の大陸において発生していたことを示唆する EBLs を同定した（図 11C）。1つ目は、マダガスカルキツネザルのゲノムにおいて特異的に検出された EBL である。この遺伝子座は、アフリカに生息する姉妹種、ショウガラゴのゲノムでは検出されなかった。したがって、この EBL の由来となった古代ボルナウウイルスの感染は、キツネザルの祖先がマダガスカル島へ移住した後に発生したと考えられた。ただし先行研究により、マダガスカルキツネザルの祖先がアフリカ大陸からマダガスカル島へ移住した時期は、約

6,000 万年前から 5,000 万年前であったと推定されているのに対して(68)、この EBL 遺伝子座の挿入年代は約 5,900 万年前から 3,700 万年前であったと推定されている。したがって、この EBL の挿入がアフリカ大陸で発生し、マダガスカル大陸に移住したキツネザル系統のみで固定化した可能性も否定できない。2 つ目は、真猿下目または広鼻小目の祖先ゲノムに挿入された EBLs である。広鼻小目の祖先は、真猿下目の祖先から種分岐する過程において、ユーラシア大陸からアフリカ大陸を経由し南アメリカ大陸へと移住したと考えられている(63, 69, 70)。したがって、真猿下目または広鼻小目の祖先ゲノムに挿入された EBLs は、霊長類祖先がこれら 3 つの大陸に移住する間に発生したボルナウイルスの感染に由来すると考えられた。

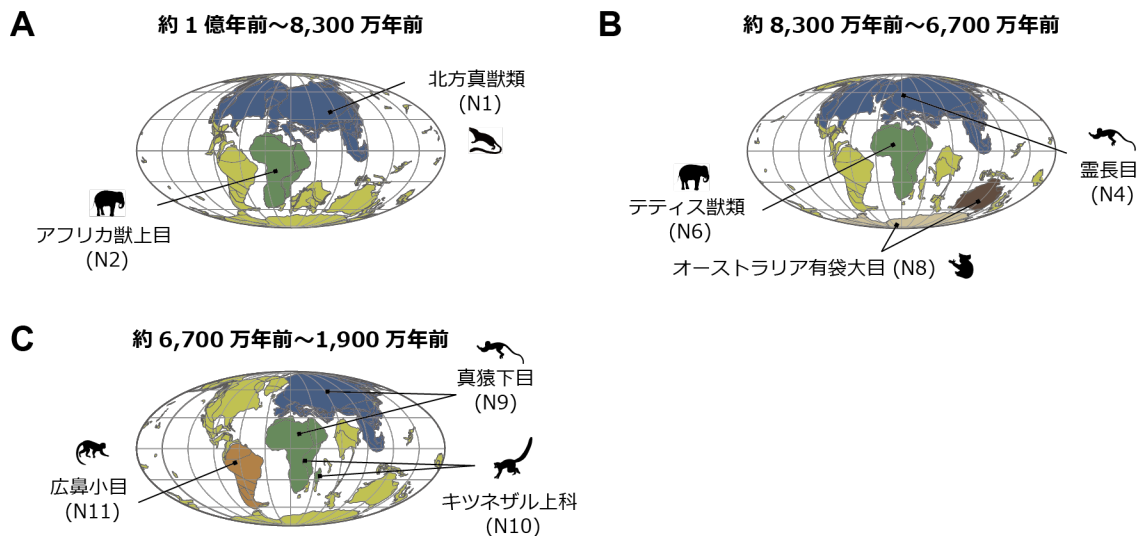


図 11. 古代ボルナウイルスの感染が発生していた地域の推定

下記の時代における祖先生物の生息域：(A) 約 1 億年前から 8,300 万年前まで、(B) 約 8,300 万年前から 6,700 万年前まで、(C) 約 6,700 万年前から 1,900 万年前まで。生物地理学についての文献(61-70)に基づき、EBLs の挿入が発生した時代に宿主生物の祖先が生息していたと考えられている大陸を矢印で示している。注釈は図 10 と対応しており、以下の生物の共通祖先を意味する：北方真獣類(N1)、アフリカ獣上目(N2)、霊長目(N4)、テティス獣類(N6)、オーストラリア有袋大目(N8)、真猿下目(N9)、キツネザル上科(N10)、広鼻小目(N11)。ボルナウイルスの感染が発生していたと考えられた大陸は、以下のとおり色分けした：ローラシア大陸またはユーラシア大陸(青)、アフリカ大陸(緑)、南極大陸(薄茶)、オーストラリア大陸(濃茶)、南アメリカ大陸(オレンジ)。それぞれの時代における大陸の位置は、Ocean Drilling Stratigraphic Network Plate Tectonic Reconstruction Service (<https://www.odsn.de/odsn/services/paleomap/paleomap.html>) から取得した。

2-5. 霊長類祖先に感染していたボルナウイルス系統の時系列変化

興味深いことに、図 10A の結果から、EBLs の挿入が霊長類の進化過程において繰り返し発生してきたことを見出した。そこで、どのような古代ボルナウイルスが霊長類祖先に感染してきたのかを明らかにするために、系統解析を行った (図 12)。その結果、霊長類のゲノムに存在する EBLNs は、少なくとも 6 つの異なる系統のボルナウイルスに由来していることが示された。例えば、カルボウイルス属の系統樹において、霊長類のゲノムに存在する EBLNs は大きく 2 つのクレードに分類された (図 12A)。このうち、クレード 1 は約 9,600 万年前に挿入された EBLNs によって、クレード 2 は約 6,700 万年前から 2,900 万年前の間に挿入された EBLNs によって構成されていた。また、オルソボルナウイルス属の系統樹においても、霊長類のゲノムに存在する EBLNs は 3 つの異なるクレードに分類され、それぞれのクレードには異なる時代に挿入された EBLNs が含まれていた (図 12B)。これらの結果は、霊長類の祖先が時代の移り変わりとともに異なる系統のボルナウイルスに感染してきたことを示唆している。

次に、霊長類の祖先に感染していたボルナウイルスの多様性を評価するために、図 12A-C の系統樹を用いて古代ボルナウイルスの間の遺伝的多様性を算出した。現代のボルナウイルス種を区別する遺伝的距離を比較基準とした解析により、霊長類祖先に感染していたクレード 1 からクレード 6 の古代ボルナウイルス系統は、現代のボルナウイルス種よりも遺伝的に多様であったことが示された (図 13)。これらの結果は、異なるボルナウイルス種に相当する、多様な系統のボルナウイルスが霊長類祖先に感染してきたことを示唆している。

2-6. 霊長類祖先とボルナウイルスとの数千万年にわたる共存

前項では、図 12 の結果から、霊長類の進化過程におけるボルナウイルスの流行の変化を明らかにした。一方で、図 12 の系統解析により、数千万年にわたって霊長類祖先に内在化を繰り返していたボルナウイルス系統が同定され、これらのボルナウイルス系統が霊長類祖先と長期的な共存関係を築いていた可能性を見出した。例えば、クレード 2 系統のカルボウイルスに由来する EBLNs は、約 6,700 万年前から 2,900 万年前までの間で、真猿下目および広鼻小目の祖先ゲノムに計 7 回挿入されたと推定されている（図 12D）。また、クレード 5 系統のオルソボルナウイルスに由来する EBLNs も、約 6,700 万年前から 1,900 万年前までの間で、真猿下目および狭鼻小目の祖先ゲノムに計 7 回挿入されている。こうした結果は、これら 2 つの系統のボルナウイルスが数千万年にわたって霊長類祖先と共存していたことを示唆している。したがって、前項「2-5. 霊長類祖先に感染していたボルナウイルス系統の時系列変化」と本項の結果より、霊長類祖先と古代ボルナウイルスの共進化過程は、1) 霊長類祖先への異なる系統のボルナウイルスの感染、および 2) 長期的なウイルスと宿主生物との共存関係の構築といった異なるイベントにより構成されていたことが明らかとなった。

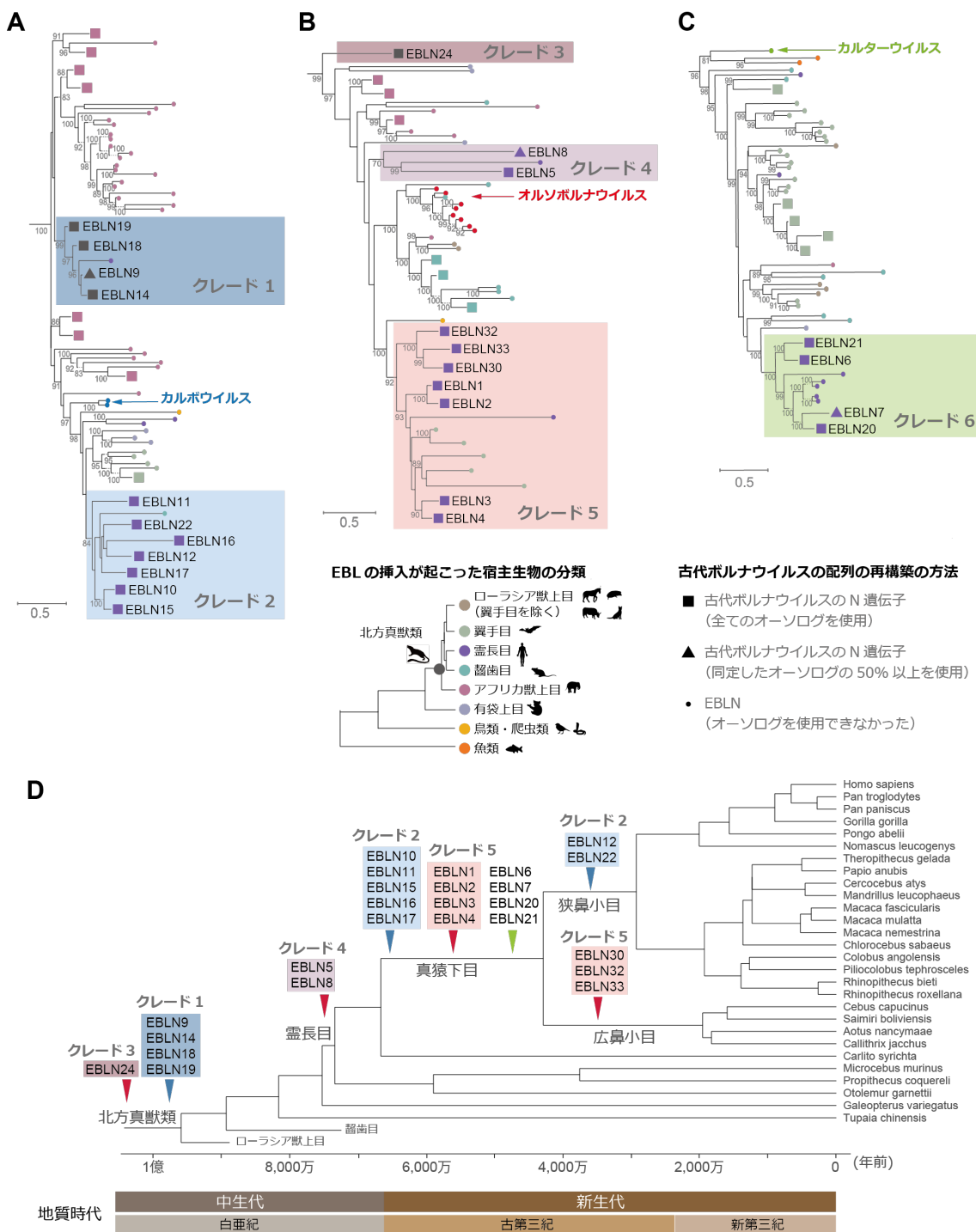


図 12. 霊長類祖先に感染していた古代ボルナウイルス

(A) カルボウイルス属、(B) オルソボルナウイルス属、(C) カルターウイルス属における古代ボルナウイルスと現代のボルナウイルスの N 遺伝子の系統関係。これらの系統樹は、EBLNs と現代のボルナウイルスの N 遺伝子のアミノ酸配列を用いて最尤法により構築した。現代のボルナウイルスは矢印で示されている（カルボウイルス：青、オルソボルナウイルス：赤、カルターウイルス：緑）。古代のボルナウイルスの N 遺伝子の配列は、オーソログを用いて再構築した（詳細は「第四章 材料と方法」を参照）。ノードの形は、古代ボルナウイルスの N 遺伝子の配列をどのようにして再構築したかを示している：本研究で同定したすべてのオーソログを使用（四角）、本研究で同定したオーソログの 50%以上を使用（三角）、系統解析にオーソログ配列を使用できなかったため EBLN 配列そのものを使用（丸）。ノードの色は EBLNs の挿入が起こった宿主生物の分類に対応している。クレード 1 からクレード 6 は、霊長類祖先に感染していたボルナウイルスの系統を示す。枝の支持率(%)はブートストラップ法により算出した。スケールバーは遺伝的距離（座位あたりの置換数）を示す。(D) 霊長類の進化過程における EBLNs の挿入の歴史。矢頭はそれぞれの EBLNs の由来となった古代ボルナウイルスの挿入時期を示している。クレード 1 からクレード 5 は (A-C) で示している系統に対応している。

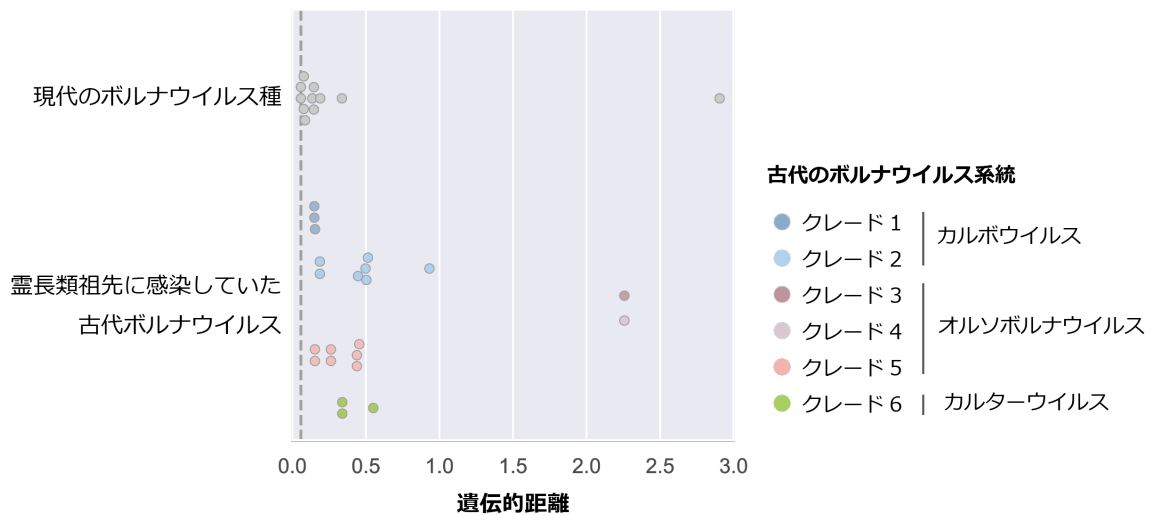


図 13. 古代のポリオウイルス間の遺伝的多様性

図 12A-C の系統樹における現代のポリオウイルス種間、または霊長類祖先に感染していた古代のポリオウイルス配列間の遺伝的距離のうち、最も近縁な配列との遺伝的距離を示す（詳細は「第四章 材料と方法」を参照）。灰色の点線は、現代のポリオウイルス種を区別する遺伝的距離を示す。霊長類祖先に感染していた古代ポリオウイルスについては、すべてのオーソログを用いて再構築された配列のみを解析に用いたため、EBLN7、EBLN8、EBLN9 のデータは本図には含まれていない。

第三章

考察

3-1. EVEs を用いた古ウイルス学研究

近年、ウイルス感染症が相次いで発生している状況をかんがみると、私たちの祖先もまた、ウイルスの脅威に繰り返し晒されてきたのではないかと予想できる。一方で、ウイルスは物理的化石を残すことがないため、過去のウイルス感染を調査することは困難であると考えられてきた。しかし、生物のゲノムには、過去に感染したウイルスに由来する遺伝子配列（EVEs）が存在することが明らかとなり、ウイルスの感染を地質学的タイムスケールにおいて調査することが可能となった(7)。これまでに、EVEs を用いた古ウイルス学研究により、ウイルスの長期的進化や、宿主生物との共進化過程が解明されてきた(7-11)。

シーケンス技術の発展に伴い、さまざまな生物種のゲノムデータが次々と解読・報告されている(<https://www.ncbi.nlm.nih.gov/refseq/statistics/>)。今後、生物のゲノムデータへのアクセシビリティがさらに向上することで、新たな EVEs が同定され、太古のウイルスについての理解がより一層深まることが期待される。一方で、ゲノムデータの解読に伴い、急速に増え続けている「ウイルスの分子化石」をいかに効率的に解析するかという点が、古ウイルス学における重要な課題の1つとなっている(7,71)。特に、EVEs の挿入年代は、過去のウイルス感染の理解に不可欠な情報であり、これまでにさまざまなアプローチによる推定方法が開発されてきた(7)。しかし、レトロウイルスに由来する EVEs 以外では、数百から数千におよぶ EVE 配列の挿入年代を一挙に推定することは困難であった。

「ボルナウイルスの分子化石」である EBLs の発見により、過去のボルナウイルスの感染をスナップショットとして捉えられるようになった(8, 14, 28)。しかし、長期にわたるボルナウイルスの感染の歴史を理解するためには、EBLs の由来となった古代ボルナウイルスの宿主生物・存在年代・多様性を統合的に解析する必要がある。そこで本研究では、大規模なゲノムデータ解析により 131 種の脊椎動物において 1,400 を超える EBL 配列を同定した(図 7)。さらに、大量の EBL 配列の挿入年代を効率的に推定するために、ネットワーク解析に基づくオーソログ推定法を新たに開発し、約 1 億年にわたるボルナウイルスの感染の歴史を再構築した(図 10)。

3-2. EBLs を用いた研究の限界

本研究では、EBLs の網羅的解析により地質学的タイムスケールにおけるボルナウイルスの感染の歴史を再構築した (図 10)。一方で、他のあらゆる化石記録と同様に、EBLs は疎に点在する不完全な化石記録であることに留意しなければならない。なぜなら、感染したすべてのボルナウイルスが宿主生物のゲノムに挿入されるわけではないためである (7, 14, 26, 71)。特にボルナウイルスは、RNA 配列の逆転写および生物ゲノムへの挿入を自律的に行うことができないため、ボルナウイルスの内在化はまれなイベントであることが予想される。また、ボルナウイルスの遺伝子配列が生物のゲノムに挿入されたとしても、長期的な進化の過程においてウイルスに由来する配列が失われる可能性もある。例えば、挿入されたボルナウイルス配列が宿主生物集団において固定化されなかった場合や、宿主生物の絶滅に伴い、その生物のゲノムに存在していたボルナウイルス配列も失われてしまう場合が考えられる。したがって、EBLs は長期的進化の過程で発生した多様なボルナウイルスのうち、ごく一部のウイルスに由来する分子化石であることに注意する必要がある。EBLs は過去のボルナウイルスの感染を調査するために欠かせない材料であるが、こうした潜在的な限界を前提とした上で、古代のウイルスの感染について慎重に議論を進める必要がある。

3-3. ボルナウイルスの長期的進化の背後にあるメカニズム

RNA ウイルスの進化に関する先行研究では、現代のウイルスを用いた系統解析により、「ボルナウイルスの長期的進化は脊椎動物との共分岐により駆動されてきた」ことを示唆している (図 14A) (37)。ウイルスと宿主生物との共分岐とは、宿主生物が種分岐する過程で、新たな生物種に適応したウイルスが出現し、結果として、宿主生物の進化に呼応するようにウイルスが進化していく現象を指す。

しかし、本研究により示された現代と古代のボルナウイルスの系統関係からは、ボルナウイルスの長期的進化をウイルスと宿主生物との共分岐のみで説明することは難しいと考えられた。この理由には、1)時代の移り変わりに伴い、流行していたボルナウイルスの系統が変化してきたこと、2)それぞれの時代に存在したボルナウイルスが多様な宿主生物に感染を広げていたことがあげられる (図 14B)。まず本研究では、霊長類の進化過程において流行していたボルナ

ウイルスの系統が変化してきたことを明らかにした（図 12）。加えて、図 14B では、霊長類祖先に感染していたボルナウイルスに近縁なウイルスが、同じ時代に他の生物にも感染を広げていた可能性を示した。例えば、クレード 1 系統のカルボウイルスに近縁なウイルスは、中生代後期（約 8,300 万年前から 6,600 万年前ごろ）においてアフリカ獣上目やテティス獣類の祖先に感染していたようである（図 14B の EBLN63、EBLN66、EBLN49、EBLN41）。また、クレード 2 系統のカルボウイルスに近縁なウイルスは、約 5,300 万年前にココウモリの祖先に感染していたと推定されている（図 14B の EBLN59）。同様の傾向はオルソボルナウイルスの系統樹でも観察されている。例えば、現代のオルソボルナウイルスは遺伝的に近縁な単一のクレードを形成するにもかかわらず、哺乳類、鳥類、爬虫類といった多様な宿主生物に感染することが報告されている(35)。このような結果は、時代の移り変わりとともに流行していたボルナウイルスの系統が変化してきたこと、およびそれぞれの時代に存在したボルナウイルスがさまざまな宿主生物に感染を広げていたことを示唆している。したがって本研究では、ボルナウイルスの長期的進化は、ウイルスと宿主との共分岐だけでなく、時代の移り変わりに伴うウイルスの流行の変化およびウイルスの異種間伝播によって構成されてきた可能性を提唱する。

ただし前項「3-2. EBLs を用いた研究の限界」のとおり、EBLs はすべてのボルナウイルスの感染を記録しているわけではないため、本研究では捉えきれないウイルスと宿主生物との関係が存在した可能性は否定できない。しかし、本研究により得られた結果は、現代のボルナウイルスを用いた解析からだけではウイルスの長期的進化の全容を解明することは困難であることを示唆している。したがって今後は、本研究において同定した「ボルナウイルスの分子化石」を活用することで、ボルナウイルスの長期的進化について、さらなる検証を行う必要がある。

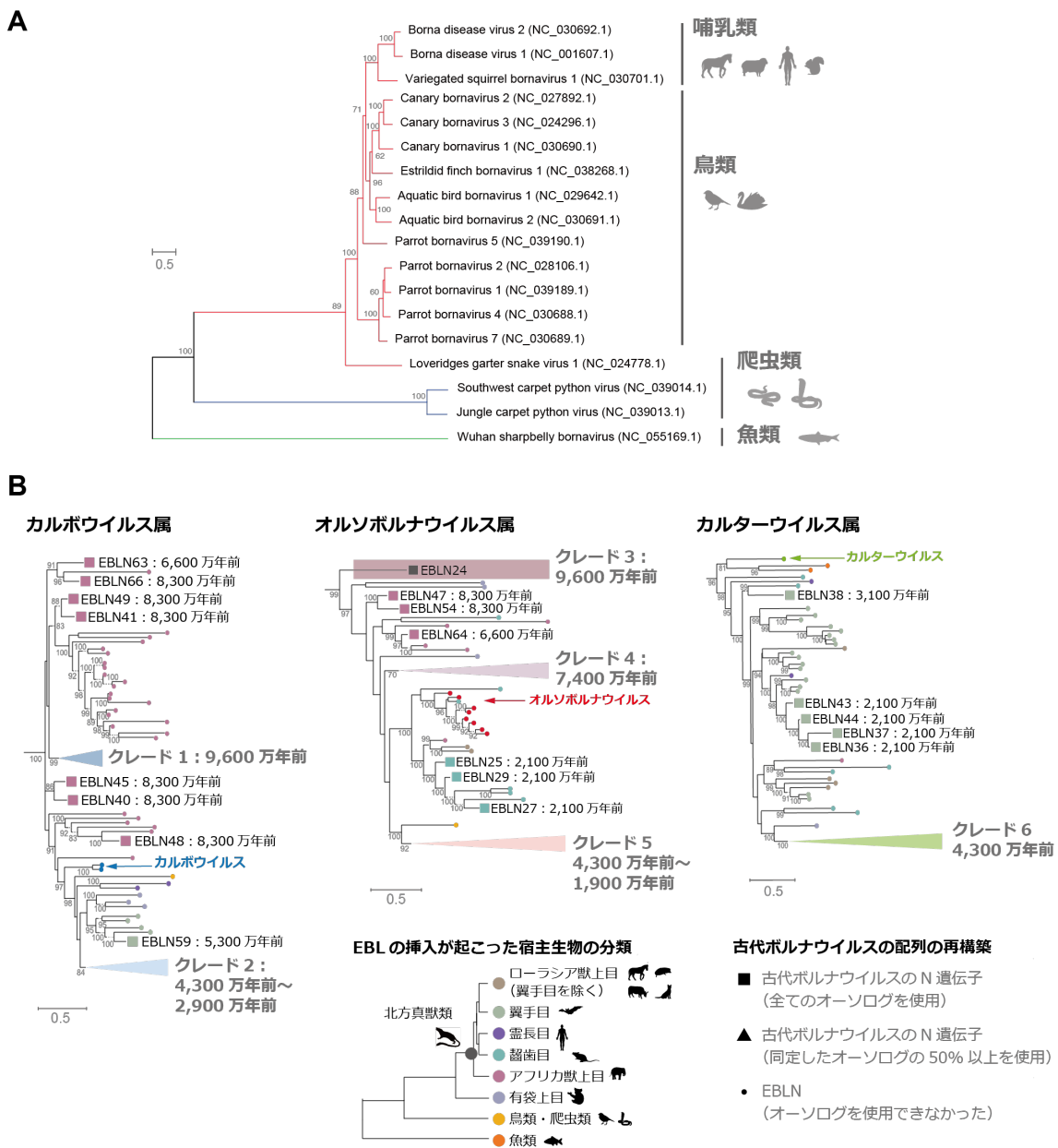


図 14. ボルナウイルスと脊椎動物との共分岐についての考察

(A) 現代のボルナウイルスの配列のみを用いて作成した系統樹。枝の色はオルソボルナウイルス (赤)、カルボウイルス (青)、カルターウイルス (緑) を示す。ブートストラップ法により枝の支持率(%)を算出した。スケールバーは遺伝的距離 (座位あたりの置換数) を示す。系統樹の右側には、それぞれのボルナウイルスの代表的な宿主生物を示す。先行研究ではこのような系統関係に基づき、「ボルナウイルスは脊椎動物との

共分岐により進化してきた」ことを示唆している(37)。(B) 古代のボルナウイルスと現代のボルナウイルスの N 遺伝子のアミノ酸配列を用いた系統樹。この系統樹は図 12A-C を改変したものである。それぞれの外部節には、EBLN 遺伝子座の名前とその遺伝子座の挿入年代が記載してある。本研究により、1)時代の移り変わりとともに、流行していたボルナウイルスの系統が変化してきたこと、2)それぞれの時代に存在したボルナウイルスが多様な宿主生物に感染を広げていたことを示唆した。したがって、ボルナウイルスの長期的進化は、ウイルスと脊椎動物との共分岐だけでなく、時代ごとのウイルスの流行の変化およびウイルスの異種間伝播によって構成されていた可能性が考えられた。

3-4. ウイルスと宿主との長期的共存が感染の拡大をもたらした可能性

図 12 の結果から、2つの系統のボルナウイルスが霊長類祖先と数千万年にわたって共存していた可能性を見出した。さらに本研究では、祖先生物の生息域についてのこれまでの知見と本研究で得られたボルナウイルスの内在化時期を統合的に解析した結果、霊長類祖先と長期的に共存していたオルソボルナウイルス系統の感染が複数の大陸において発生していたことを示唆した。まず、図 12D に示すように、クレード 5 系統のオルソボルナウイルスは、真猿下目の祖先および広鼻小目の祖先において感染と内在化を繰り返していたと推定された。霊長類祖先とクレード 5 系統のボルナウイルスが共存していたと考えられた期間において、真猿下目の祖先はユーラシア大陸またはアフリカ大陸に生息していたのに対して、広鼻小目の祖先は、真猿下目から種分岐する過程で、アフリカ大陸から南アメリカ大陸へと移住したと考えられている (図 11C) (63, 69, 70)。したがって、クレード 5 系統のオルソボルナウイルスは、霊長類祖先の移住に伴い、異なる大陸に感染を広げた可能性が考えられた。

加えて、クレード 5 系統のオルソボルナウイルスに由来する EBLNs は、キクガシラコウモリ、チスイコウモリ、ユビナガコウモリのゲノムにおいても同定された (図 15A)。生物地理学における先行研究により、これらの EBLNs の挿入時期において、キクガシラコウモリの祖先はユーラシア大陸に、チスイコウモリの祖先は南アメリカ大陸に、ユビナガコウモリの祖先は複数の異なる大陸に生息していたと推定されている (図 15B) (72)。これらの結果を統合すると、クレード 5 系統のオルソボルナウイルスは霊長類祖先の移住とともに大陸間を移動し、複数の祖先へと感染を広げた可能性が考えられた。以上より、ウイルスと宿主生物の長期的な共存関係が、ウイルスの地理的分布の拡大、さらには新たな生物へのウイルス伝播をもたらしたことを示唆した。

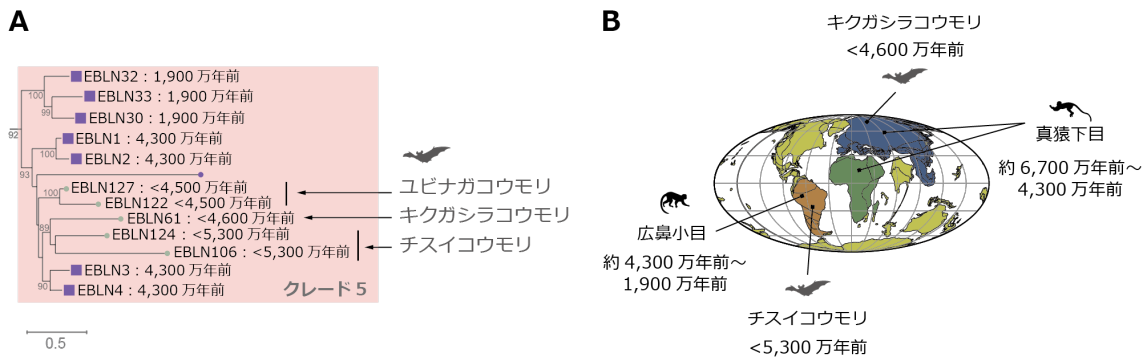


図 15. ウイルスと宿主との長期的共存による感染拡大についての考察

(A) コウモリゲノムで同定されたクレード 5 系統に属する EBLNs。この系統樹は図 12B の一部を拡大したものである。それぞれのノードには、EBLN 遺伝子座の名前とその遺伝子座の挿入年代が記載してある。ノードの色は EBLNs の挿入が起こった宿主生物の分類を示す（真猿下目または広鼻小目：紫、翼手目：薄緑）。ブートストラップ法により枝の支持率(%)を算出した。スケールバーは遺伝的距離（座位あたりの置換数）を示す。この系統樹により、クレード 5 系統のオルソボルナウイルスは、霊長類祖先だけでなくコウモリ祖先にも感染していた可能性が示された。(B) クレード 5 系統のオルソボルナウイルスに感染していたと考えられる宿主生物の生息域。本図では、約 6,700 万年前から 1,900 万年前における真猿下目と広鼻小目の祖先(63, 69, 70)、およびコウモリ祖先の生息域(72)を示す。ユビナガコウモリの祖先は複数の大陸に生息していたと推定されていたため、本図には含まれていない。

3-5. 宿主生物ごとのボルナウイルスの内在化の頻度の違い

本研究では、脊椎動物の進化過程におけるボルナウイルスの内在化の時期と数を解析し、これらの傾向が宿主系統によって大きく異なっていることを明らかにした(図 10)。例えば、鳥類や爬虫類においては、ボルナウイルスの内在化は比較的まれな現象であったことが示され、こうした結果は先行研究での報告と一致していた(29, 73)。さらに本研究では、ボルナウイルスの内在化の頻度が哺乳類の系統間においても異なっていることを示した。例えば、北方真獣類の子孫系統に着目すると、霊長目や翼手目、齧歯目ではボルナウイルスの内在化が繰り返し発生していたのに対して、翼手目を除くローラシア獣類ではほとんど発生していなかったことが明らかとなった。

ボルナウイルスの内在化の頻度は、ウイルスの感染、生殖細胞ゲノムへのウイルス配列の挿入、および挿入されたウイルス配列の宿主子孫への遺伝といった複数のプロセスによって決定されると考えられる。本研究では、過去に発生したボルナウイルスの感染について主に議論したが、ボルナウイルス配列の挿入効率や子孫生物への遺伝についての検証も、生物種間におけるボルナウイルスの内在化の頻度の違いを理解するために不可欠である。例えば、RNA ウイルスの遺伝子配列を生物のゲノムに挿入すると考えられているレトロトランスポゾンの活性は、ボルナウイルスの挿入効率に影響を与えると予想される(74, 75)。また、挿入されたボルナウイルス配列に対する選択圧や子孫生物の集団サイズの変動は、ウイルス配列の固定化の頻度を変化させうるだろう(76)。生物系統ごとの EBLs の内在化の頻度の違いと、ウイルスの感染規模、ウイルス配列の挿入効率、子孫生物への遺伝の様式といったそれぞれのプロセスとの関連については、集団遺伝学に基づくシミュレーション解析等により、さらなる検証が必要である。

3-6. EVEs の活用によるウイルスの多様性の解明

EVEs の同定方法の改良は、古ウイルス学における重要な課題の 1 つである (26, 71, 77, 78)。EVEs の同定は、現代のウイルスの遺伝子配列に類似した領域を生物ゲノムにおいて探索することで行われる。したがって、EVEs と現代のウイルスとの配列類似性は、EVEs の検出感度に大きな影響を与える。本研究では、カルボウイルスやカルターウイルスといった新たなボルナウイルスの発見をきっかけに、EBL 探索に使用するクエリ配列を充実させることで、これまでの探索では見逃されていた EBLs を大量に同定することに成功した (図 7)。この結果は、現代のウイルスの多様性の解明が古代のウイルスの多様性の解明に直結していることを意味する。

同時に、過去のウイルスの遺伝子配列である EVEs を、現代のウイルスの多様性の解明に活用することも可能であると考えられる。なぜなら、現代のウイルス配列の同定方法も、EVEs の探索方法と同様に、既知のウイルスとの配列類似性に依存しているためである (79, 80)。さらに本研究において同定された EBLNs は、現代のボルナウイルスとは全く異なる系統に属する、多様なボルナウイルスに由来する配列であることが示されている (図 6)。したがって、EBLs の配列情報を現代のウイルスの探索に利活用することで、既知のボルナウイルスとは異なる系統に属する、新たなボルナウイルスが発見される可能性がある。以上より、現代のウイルスと古代のウイルスの配列情報を互いにフィードバックすることで、ウイルスの多様性と進化の理解が促されることが考えられる。

第四章

材料と方法

4-1. 真核生物のゲノムデータにおける EBLs の探索

まず、2019年4月時点にアメリカ国立生物化学情報センター (NCBI) が RefSeq genomic database(81)として公開していた 969 種の真核生物のゲノムデータにおいて、ボルナウイルスに由来する遺伝子配列を tBLASTn (version 2.6.0+) (82) により探索した。本探索では、ボルナウイルス科に分類されるウイルス配列をクエリとして使用し (表 1)、E-value が 0.01 以下で検出された配列を以降の解析に使用した。

EBLs の多くは、フレームシフト変異により断片化された配列として検出されていたため、宿主ゲノムにおける位置、および現代のボルナウイルスタンパク質との配列アライメントにおける位置に基づいて、ボルナウイルス様配列を結合し、EBL 配列を再構築した (図 1A)。2 つ以上のボルナウイルス様配列が同じゲノム位置で検出された場合には、tBLASTn においてより低い E-value で検出された配列を信頼性の高い配列とみなし、EBL 配列の再構築に使用した。現代のボルナウイルスのタンパク質と tBLASTn で検出されたボルナウイルス様配列のアミノ酸配列を用いた多重配列アライメント (MSA) は MAFFT (version 7.427) (83)により構築した。

さらに、再構築した配列がボルナウイルスに由来する配列であることを確認するために、NCBI Refseq protein database として公開されていたタンパク質配列と、再構築した配列との類似性を BLASTP (version 2.9.0+) により算出した。再構築した配列が、ボルナウイルス以外のウイルスタンパク質または宿主タンパク質に対して最も高い類似性を示した場合、その配列はボルナウイルスに由来するものではないと判定し、以降の解析から除外した。ここまでの解析で得られた 1,466 配列のうち、1 配列は無脊椎動物のゲノムに存在していた。本研究では脊椎動物の進化過程におけるボルナウイルスの感染に焦点を当てるため、この配列を以降の解析から除外した。

表 1. EBLs の探索においてクエリとして使用したボルナウイルス配列

| INSD アクセッション 番号 | ウイルス属 | ウイルス名 | ウイルス タンパク 質 |
|-----------------------|------------|-------------------------------|-------------------|
| CAC70640.1 | オルソボルナウイルス | Borna disease virus 1 | N |
| CAC70641.1 | オルソボルナウイルス | Borna disease virus 1 | X |
| CAC70642.1 | オルソボルナウイルス | Borna disease virus 1 | P |
| CAC70643.1 | オルソボルナウイルス | Borna disease virus 1 | M |
| CAC70644.1 | オルソボルナウイルス | Borna disease virus 1 | G |
| CAC70645.1 | オルソボルナウイルス | Borna disease virus 1 | L |
| ATY47614.1 | カルボウイルス | Jungle carpet python virus | N |
| ATY47617.1 | カルボウイルス | Jungle carpet python virus | X |
| ATY47615.1 | カルボウイルス | Jungle carpet python virus | P |
| ATY47613.1 | カルボウイルス | Jungle carpet python virus | G |
| ATY47616.1 | カルボウイルス | Jungle carpet python virus | M |
| ATY47612.1 | カルボウイルス | Jungle carpet python virus | L |
| ATY47620.1 | カルボウイルス | Southwest carpet python virus | N |
| ATY47623.1 | カルボウイルス | Southwest carpet python virus | X |
| ATY47621.1 | カルボウイルス | Southwest carpet python virus | P |
| ATY47619.1 | カルボウイルス | Southwest carpet python virus | G |
| ATY47622.1 | カルボウイルス | Southwest carpet python virus | M |
| ATY47618.1 | カルボウイルス | Southwest carpet python virus | L |
| AVM87536.1 | カルターウイルス | Wuhan sharpbelly bornavirus | N |
| AVM87537.1 | カルターウイルス | Wuhan sharpbelly bornavirus | X |
| AVM87538.1 | カルターウイルス | Wuhan sharpbelly bornavirus | P |
| AVM87539.1 | カルターウイルス | Wuhan sharpbelly bornavirus | G |
| AVM87540.1 | カルターウイルス | Wuhan sharpbelly bornavirus | M |
| AVM87541.1 | カルターウイルス | Wuhan sharpbelly bornavirus | L |

4-2. EBLs の挿入年代の推定

EBLs の挿入年代は、遺伝子オーソロジー解析により推定した(7)。本研究では1,400 を超える EBLs のオーソログ関係を決定するために、EBLs の挿入部位の配列類似性に基づいたネットワーク解析を行い、オーソロググループを抽出した(図8)。

まず、EBLs の挿入部位の塩基配列を seqkit (version 1.5.0)(84)により抽出した。EBLs の挿入部位に含まれる反復配列の除去は、RepBase RepeatMaker ライブラリ(85)を用いて RepeatMasker (version open-4.0.9)により行った。挿入部位の配列として、EBLs の場合は挿入部位の上流下流それぞれ 15.0kb、それ以外の EBLs は上流下流それぞれ 10.0kb を抽出した。EBL 挿入部位の配列類似性を定義するために、BLASTN (version 2.9.0+)によりペアワイズアライメントを作成し、それぞれの配列間におけるアライメントカバレッジ(クエリ配列長あたりのアライメントされたサイトの割合)を算出した。次に、配列類似性ネットワークを EBL 挿入部位間のアライメントカバレッジに基づいて作成し、コミュニティ構造の抽出を行った。今回は、アライメントカバレッジが 9.0%以上であった EBL 配列をエッジで結合することによりネットワークを構築した。配列類似性ネットワークの構築についての条件検討は、次項「4-3. ネットワーク解析による年代推定法の検証」において詳述する。ネットワーク解析には、networkx (version 2.4)を用い、コミュニティ構造の抽出は Louvain 法により実施した(86)。最後に、TimeTree データベース(87)が提供する脊椎動物の種分岐の年代に従って、EBLs の挿入時期を決定した。

遺伝子オーソロジー解析による EBLs の挿入年代の推定結果は、以下のパターンに分類される(図16)。このうち、挿入年代について誤った推定結果を導く可能性として、オーソログな配列が検出できないために挿入年代を過小推定する場合(図16C)と、オーソログな関係にない配列のコンタミネーションにより挿入年代を過大推定する場合(図16D)が考えられる。

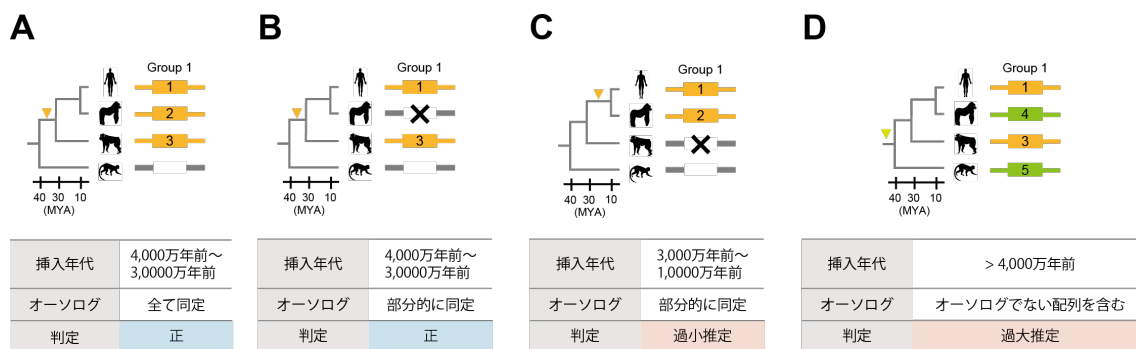


図 16. オーソログの同定と挿入年代の推定の関係

それぞれのパネルにおける系統樹は、宿主生物の進化過程を示している。本図では、(A)に示されるような、配列 1、2、3 がオーソログな関係にあり、これらの配列は約 4,000 万年前から 3,000 万年前に発生した挿入イベントに由来するという場合を考える。ネットワーク解析において推定された挿入年代と同定されたオーソログの組み合わせにより、結果は以下の 4 つに分類される：(A) すべてのオーソログが同定され、挿入年代が正しく推定された、(B) 部分的にしかオーソログが同定されなかったが、挿入年代は正しく推定された、(C) 部分的にしかオーソログを同定できなかったため、挿入年代を過小推定した、(D) オーソログでない配列をオーソログと誤認定してしまったため、挿入年代を過大推定した。

本研究では、ネットワーク解析による EBLs の挿入年代の推定の正確性を向上させるために、宿主生物の系統関係およびアライメントカバレッジを確認した (図 9)。挿入年代の過大評価を避けるためには、オーソログな関係にならない配列を同一のグループに混同することを防ぐ必要がある (図 16D)。そのため、同じ種のゲノムに存在する EBLs では、挿入部位間において 9% 以上のアライメントカバレッジが認められたとしても、それが 50% 以下であった場合は異なるオーソロググループに分類した。また挿入年代の過小評価は、より古くに挿入された遺伝子座で起こりやすいと考えられる (図 16C)。これは、ウイルス配列が挿入された後の時間経過に伴い、突然変異やゲノムリアレンジメントが発

生することで、挿入部位間における配列類似性を検出することが困難になると予想されるためである。したがって、ネットワーク解析で分離されたグループの間において低い配列類似性が検出された場合には、こうしたグループに含まれる EBLs の配列アライメントを AliTV (version 1.6.0) (88-90)により作成し、これらのグループが同一のオーソロググループに属する可能性を検証した(図 17)。結果、図 9 において白色列で示す 4 つのオーソロググループ (EBLG2、EBLL2、EBLL35、EBLL36) を再編成した。EBLG2、EBLL2、EBLL35 については、ネットワーク解析で別グループに分類されていた配列間において、70%以上の配列同一性が 40%以上の配列長で検出されたため、同一のオーソロググループに属すると判定した(図 17)。先行研究では、「EBLG2 は約 7,700 万年前にローラシア獣類の祖先ゲノムに挿入された遺伝子座である」と報告していたが(36)、本解析により北方真獣類の祖先で挿入された配列であることが示された(図 17A、D)。また EBLL36 グループには、特定のゲノム位置にタンデムに存在する EBLs が含まれている。本研究では、これらの EBLs が独立した挿入により発生したのか、挿入後の遺伝子重複により発生したのかを区別することができなかった。したがって今回は、挿入イベント数の過大評価を避けるため、これらの EBLs は同一の挿入イベントに由来したと判定し、単一のオーソロググループに分類した。

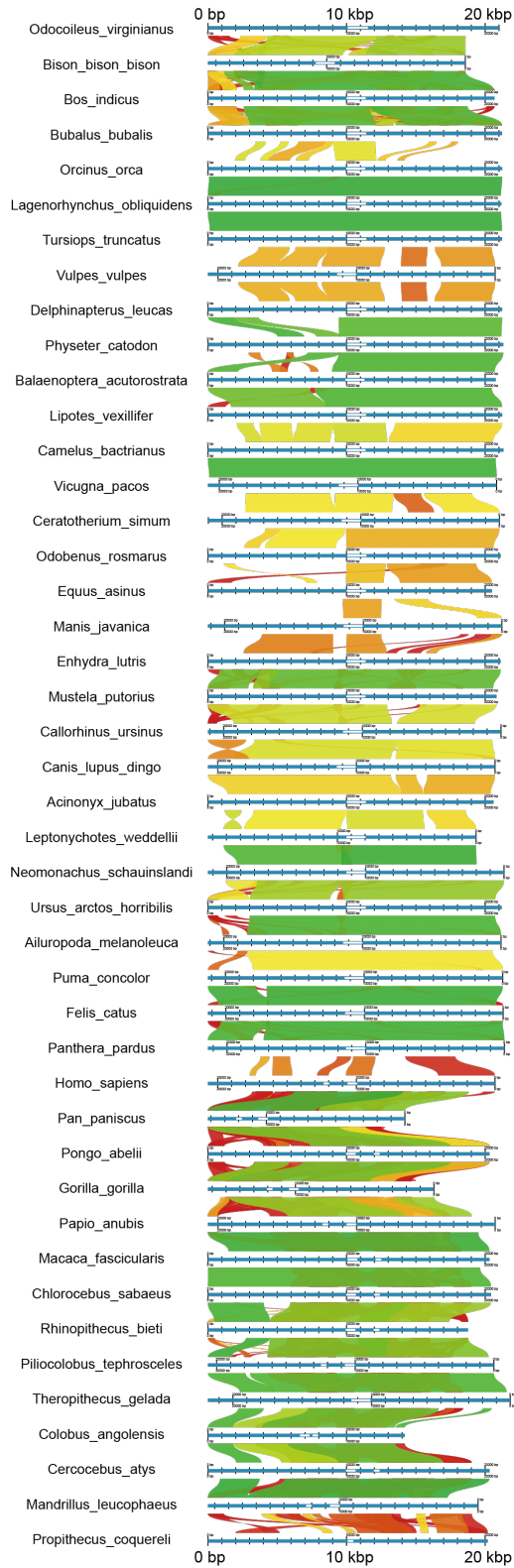
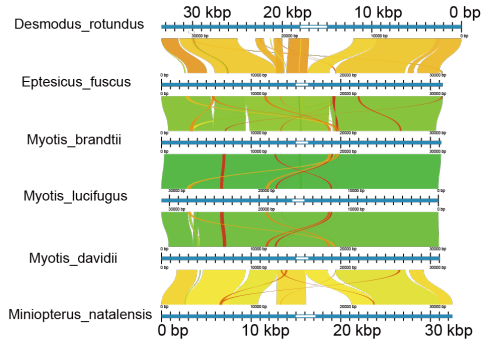
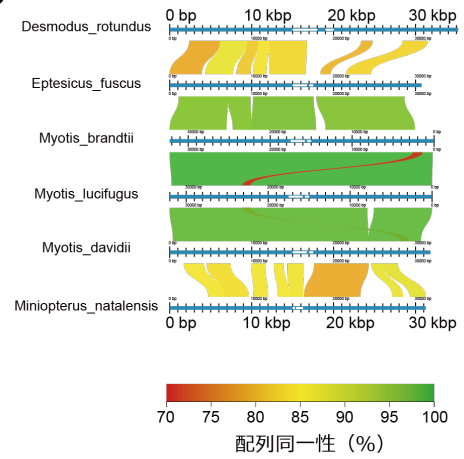
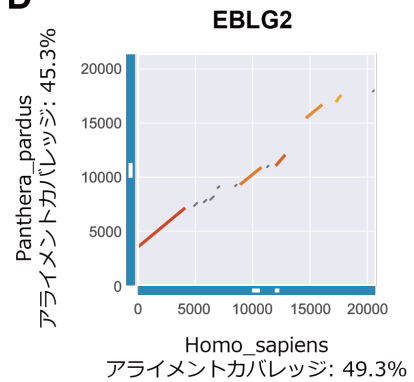
A**B****C****D**

図 17. EBLs の挿入部位間のアライメントクオリティの確認

EBLs の挿入部位の配列アライメントの模式図: (A) EBLG2、(B) EBLL2、(C) EBLL35。青いラインは EBL 挿入部位の配列を示し、その中の白いラインは EBL 配列の位置を示す。配列間のハイライトの色はアライメントにおける配列同一性を示しており、赤から緑のカラーパネルに対応している。(D) EBLG2 のオーソログ配列を用いたドットプロット。X 軸はヒトゲノムに存在する EBLG2 の周辺配列を、Y 軸はヒョウゲノムに存在する EBLG2 の周辺配列を表す。青いラインは EBLG2 の周辺配列を示し、その中の白いラインは EBLG2 の位置を示す。ドットプロットの色は (A-C) で示している配列同一性のカラーマップに対応しており、灰色のラインは 400bp 以下の短い配列がアライメントされた領域を示す。

4-3. ネットワーク解析による年代推定法の検証

ネットワーク解析による年代推定法の精度を検証するために、ゲノムアライメントデータを用いた推定方法との比較を行った。今回は、ヒトゲノムに存在するTEsの挿入年代を推定することで、2つの方法の結果を比較した(図18)。まず、ゲノムアライメントデータを用いた方法により、ヒトゲノムに存在するすべてのTEsの挿入年代の推定を行った。具体的には、LiftOver (version 372)(91)を用いて、それぞれのTE遺伝子座のオーソログを17種の哺乳類ゲノムにおいて同定し、オーソログの有無に基づき、TEsの挿入年代を推定した(図18Aの方法1)。ヒトゲノムに存在するTEsの挿入部位はRebase Repeatmaskerデータベースから取得した。ヒトと17種の哺乳類とのゲノムアライメントデータは、UCSC Genome Browserにおいて提供されていたものを使用した(92)。次に、ネットワーク解析による推定の精度をそれぞれの時代ごとに検証するために、図18Aの方法1で推定されたTEsの挿入年代に基づき、図18Cに示す9つのタイムスケールごとに100のTE遺伝子座をランダムに抽出したデータセットを作成した。ネットワーク解析による年代推定は、前項「4-2. EBLsの挿入年代の推定」に従い、それぞれのデータセットを用いて実施した(図18Aの方法2)。ただし、今回はUCSC Genome Browserによってあらかじめ反復配列がマスクされたゲノムデータを用いて、TE挿入部位間のアライメントカバレッジを算出した(<http://genome.ucsc.edu/goldenPath/credits.html>)。最後に、検出されたオーソログのセット、および推定された挿入年代を比較することで、2つの方法の一致率を算出した(図18B)。

また、配列類似性ネットワークの構築条件が年代推定の結果に与える影響を検証するために、異なるアライメントカバレッジをノードの結合基準として使用した(図18Aの方法2)。その結果、アライメントカバレッジが9.0%を超えるノード間を結合したネットワークを用いた推定が、ゲノムアライメントによる推定と最も高い一致率を示した(図18D)。このときの2つの方法の一致率は、新生代に挿入されたTEsにおいて88.0%から100.0%、中生代に挿入されたTEsにおいて43.0%から66.0%であった。

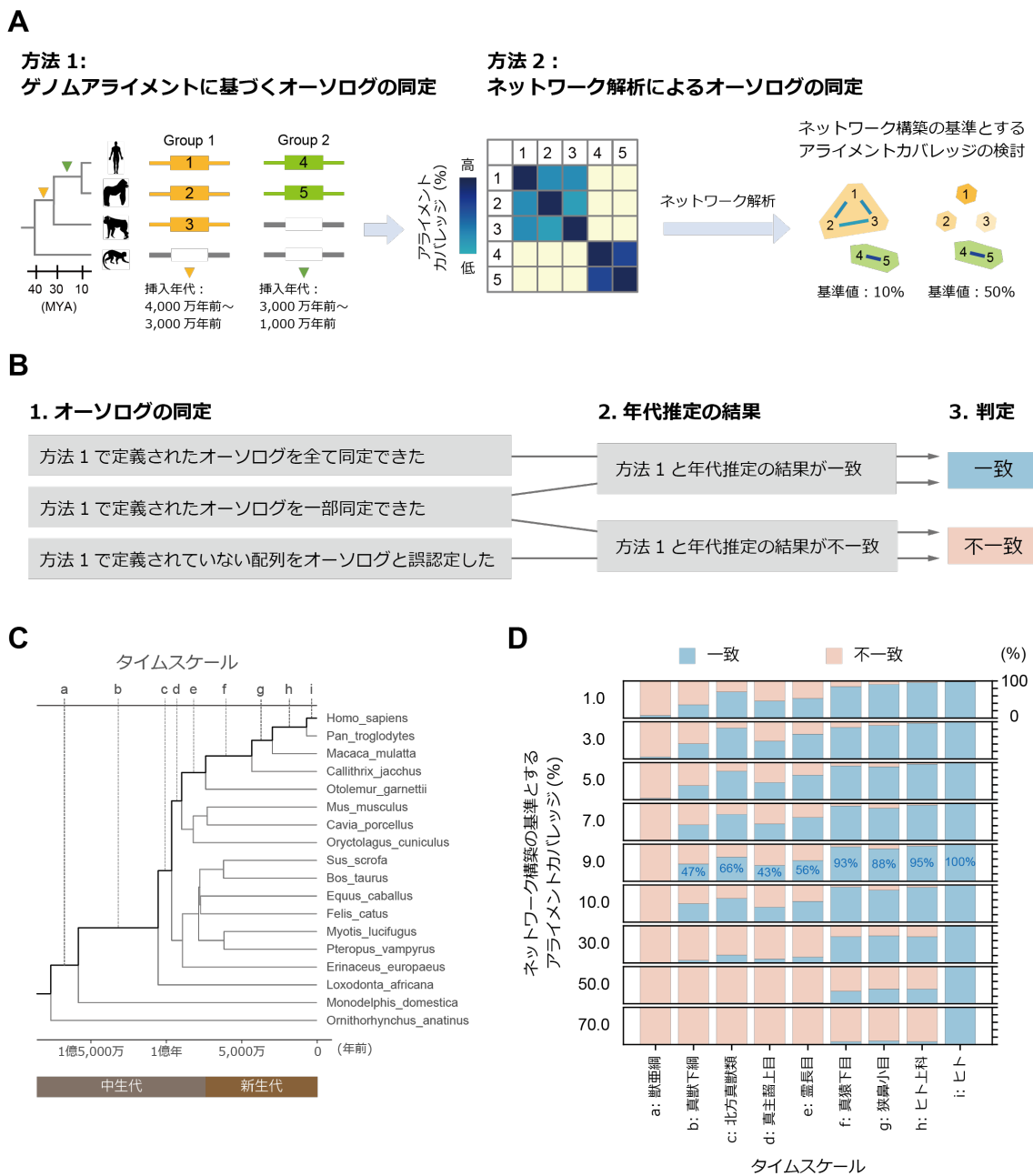


図 18. ネットワーク解析による年代推定法の検証

(A) 2つのオーソログ同定方法。方法 1 では、公共データベースですでに公開されていた、ヒトとその他哺乳類とのゲノムアライメントデータを用いてオーソログを決定した。方法 2 では、図 8 に示すネットワーク解析によりオーソログを同定した。また、ネットワークの構築条件を検討するために、異なるアライメントカバレッジを基準値とし

て用いた年代推定を行った。(B) 2つのオーソログ同定方法の結果の比較。まず、2つの方法で同定されたオーソログのセットが一致しているかを確認する。次に、オーソログの有無に基づいて推定された年代が一致しているかを確認する(オーソログの同定と年代推定の関係性については、図 16 にも記述)。(C) オーソログの同定に使用した 18 種の哺乳類の系統関係。年代推定の精度をそれぞれの時代ごとに検証するため、a から i の 9 つのタイムスケールにおいて 100 の TE 遺伝子座をランダムに抽出したデータセットを作成した。(D) 2 つの推定方法の一致率。結果が一致しているかどうかは (B) に示す方法で判定した。それぞれのパネルは、左側に示すアライメントカバレッジをネットワーク構築の基準として用いたときの 2 つの方法の一致率を示している。基準値としたアライメントカバレッジ以上の配列類似性が TEs の挿入部位間において検出された場合に、それらの配列をエッジで結合することで、ネットワークを構築した。X 軸は (C) で示したタイムスケールに対応しており、Y 軸は 2 つの推定方法の一致率を示している。本研究において EBLs の挿入年代の推定に使用した基準値(アライメントカバレッジ: 9.0%) における 2 つの方法の一致率を青字で示している。

4-4. 系統解析

EBLs と現代のウイルスタンパク質を用いた系統解析では、アミノ酸の配列長が 200 を超える EBLLs、または 100 を超える EBLNs、EBLGs を使用した。MSAs は MAFFT により構築した。図 6 に示す系統樹の構築に使用した MSAs は、30% 以上の配列がギャップであるサイトを除去し、全アライメントサイトの 70% 未満の長さしかない配列を除去することでトリミングした。また、図 12 A-C に示す EBLNs の系統樹の構築に使用した MSA は、20% 以上の配列がギャップであるサイトを除去し、全アライメントサイトの 80% 未満の長さしかない配列を除去することでトリミングした。図 14A に示す系統樹は、現代のボルナウイルスのゲノム配列を MAFFT でアライメントした MSA を用いて作成した。

すべての系統樹は、IQ-TREE (version 1.6.12)(93)を用いて、最尤法により構築した。系統樹の構築における置換モデルは、Model Finder(94)により算出されたベイズ情報量規準に基づき、以下のとおり選出した：VT+F+G4 (図 6：EBLNs)、VT+F+G4 (図 6：EBLGs)、VT+F+R3 (図 6：EBLLs)、JTT+F+G4 (図 12 A-C)、GTR+F+R4 (図 14A)。トポロジーの信頼性は、UFBoot2(95)により 1,000 回のリサンプリングにおいて算出されたブートストラップ値に基づき評価した。系統樹の図示には ggtree (version 2.2.1)(96)、または ETE3 (version 0.3)(97)を使用した。系統解析に使用した現代のウイルスの配列は表 2 に示す。

表 2. 系統解析に使用した現代のウイルスの配列

| RefSeq アクセッション番号 | ウイルス名 |
|------------------|----------------------------------|
| NC_001607.1 | Borna disease virus 1 |
| NC_030692.1 | Borna disease virus 2 |
| NC_030701.1 | Variegated squirrel bornavirus 1 |
| NC_038268.1 | Estrildid finch bornavirus 1 |
| NC_030690.1 | Canary bornavirus 1 |
| NC_027892.1 | Canary bornavirus 2 |
| NC_024296.1 | Canary bornavirus 3 |
| NC_029642.1 | Aquatic bird bornavirus 1 |
| NC_030691.1 | Aquatic bird bornavirus 2 |
| NC_039189.1 | Parrot bornavirus 1 |
| NC_028106.1 | Parrot bornavirus 2 |
| NC_030689.1 | Parrot bornavirus 4 |
| NC_039190.1 | Parrot bornavirus 5 |
| NC_030689.1 | Parrot bornavirus 7 |
| NC_024778.1 | Loveridges garter snake virus 1 |
| NC_039013.1 | Jungle carpet python virus |
| NC_039014.1 | Southwest carpet python virus |
| MG599939.1 | Wuhan sharpbelly bornavirus |
| NC_012702.1 | Midway virus |
| NC_031275.1 | Wenzhou Crab Virus 1 |
| NC_024376.1 | Sierra Nevada virus |

4-5. EBLs の由来となった古代ボルナウイルスの分類

同定した 1,465 の EBL 配列のうち 692 配列(624 EBLNs、55 EBLGs、13EBLLs) は、由来となったウイルス属を系統解析により決定できた(図 6)。系統解析に使用できなかった EBLs をウイルス属レベルに分類するために、EBLs と現代のボルナウイルスとの配列類似性の活用を試みた。まず配列類似性に基づく分類の性能を検証するために、系統解析に基づく EBLs の分類との比較を行った。配列類似性による分類では、tBLASTn において EBLs を最も低い E-value で検出した現代のボルナウイルス配列に基づき、EBLs の由来となったウイルス属を定義した。系統解析と配列類似性による分類の一致率は、EBLNs では 99.7%、EBLGs では 100%、EBLLs では 100%であった。ただし、EBLPs および EBLMs は系統解析に使用できるアライメントサイトの数が少なく、信頼できる系統樹の作成が困難であったため、これらの EBLs については上記のような分類性能の検証は行っていない。

4-6. 古代ボルナウイルスの遺伝的多様性の評価

霊長類祖先に感染していた古代ボルナウイルスの多様性を推定するために、古代のボルナウイルスと現代のボルナウイルスの N 遺伝子の遺伝的距離を算出・比較した。まず、図 12 A-C で示す系統樹においてオーソロググループごとに EBLNs の祖先配列を推定し、これを由来となった古代ボルナウイルスの N 遺伝子の配列と定義した。オーソログを使用した古代ボルナウイルスの N 遺伝子の配列の再構築は、EBLNs が挿入された後に蓄積された変異により、遺伝的多様性を過大評価しないために行った。次に、古代ボルナウイルスの配列多様性を評価するための基準として、現代のボルナウイルス種間の遺伝的距離を算出した。この評価基準を用いて、霊長類に感染していた古代ボルナウイルス間の遺伝的距離を評価した(図 13)。系統樹内のノード間における遺伝的距離の算出には ETE3 を用いた。

ただし、この解析による古代ボルナウイルスの多様性の評価は、N 遺伝子における多様性のみを用いた代替的な方法であることに留意する必要がある。ICTV による現代のボルナウイルスの分類では、完全長なウイルスゲノムの配列類似性や宿主域の違いに基づき、ウイルス種を定義している(35)。

4-7. データの公開

本研究に関連するデータ、プログラム、バイオインフォマティクスツールの詳細は、https://github.com/Junna-Kawasaki/EBL_2020において公開されている。

引用文献

1. M. A. Spyrou, K. I. Bos, A. Herbig, J. Krause, Ancient pathogen genomics as an emerging tool for infectious disease research. *Nature Reviews Genetics* **20**, 323-340 (2019).
2. A. Düx *et al.*, Measles virus and rinderpest virus divergence dated to the sixth century BCE. *Science* **368**, 1367-1370 (2020).
3. J. K. Taubenberger, Initial Genetic Characterization of the 1918 "Spanish" Influenza Virus. *Science* **275**, 1793-1796 (1997).
4. B. Mühlemann *et al.*, Ancient human parvovirus B19 in Eurasia reveals its long-term association with humans. *Proceedings of the National Academy of Sciences* **115**, 7557-7562 (2018).
5. A. T. Duggan *et al.*, 17 th Century Variola Virus Reveals the Recent History of Smallpox. *Current Biology* **26**, 3407-3412 (2016).
6. B. Mühlemann *et al.*, Ancient hepatitis B viruses from the Bronze Age to the Medieval period. *Nature* **557**, 418-423 (2018).
7. P. Aiewsakun, A. Katzourakis, Endogenous viruses: Connecting recent and ancient viral evolution. *Virology* **479-480**, 26-37 (2015).
8. A. Katzourakis, R. J. Gifford, Endogenous Viral Elements in Animal Genomes. *PLoS Genetics* **6**, e1001191 (2010).
9. A. Hayward, C. K. Cornwallis, P. Jern, Pan-vertebrate comparative genomics unmask retrovirus macroevolution. *Proceedings of the National Academy of Sciences* **112**, 464-469 (2015).
10. P. Aiewsakun, A. Katzourakis, Marine origin of retroviruses in the early Palaeozoic Era. *Nature Communications* **8**, 13954 (2017).
11. R. J. Gifford *et al.*, A transitional endogenous lentivirus from the genome of a basal primate and implications for lentivirus evolution. *Proceedings of the National Academy of Sciences* **105**, 20362-20367 (2008).
12. E. V. Koonin, M. Krupovic, V. I. Agol, The Baltimore Classification of Viruses 50 Years Later: How Does It Stand in the Light of Virus Evolution? *Microbiology and Molecular Biology Reviews* **85**, e00053-00021 (2021).
13. K. Kryukov, M. T. Ueda, T. Imanishi, S. Nakagawa, Systematic survey of non-retroviral virus-like elements in eukaryotic genomes. *Virus Research* **262**, 30-36 (2019).
14. M. Horie *et al.*, Endogenous non-retroviral RNA virus elements in mammalian genomes. *Nature* **463**, 84-87 (2010).

15. R. A. Weiss, The discovery of endogenous retroviruses. *Retrovirology* **3**, 67 (2006).
16. Human Genome Sequencing Consortium, Initial sequencing and analysis of the human genome. *Nature* **409**, 860-921 (2001).
17. Mouse Genome Sequencing Consortium, Initial sequencing and comparative analysis of the mouse genome. *Nature* **420**, 520-562 (2002).
18. J. G. N. Barreat, A. Katzourakis, Paleovirology of the DNA viruses of eukaryotes. *Trends in Microbiology* **S0966-842X**, 00164-00165 (2021).
19. E. Henckaerts *et al.*, Site-specific integration of adeno-associated virus involves partial duplication of the target locus. *Proceedings of the National Academy of Sciences* **106**, 7571-7576 (2009).
20. S. Daya, N. Cortez, K. I. Berns, Adeno-associated virus site-specific integration is mediated by proteins of the nonhomologous end-joining pathway. *Journal of Virology* **83**, 11655-11664 (2009).
21. G. Aimola, G. Beythien, A. Aswad, B. B. Kaufer, Current understanding of human herpesvirus 6 (HHV-6) chromosomal integration. *Antiviral Research* **176**, 104720 (2020).
22. C. A. Bill, J. Summers, Genomic DNA double-strand breaks are targets for hepadnaviral DNA integration. *Proceedings of the National Academy of Sciences* **101**, 11135-11140 (2004).
23. C. Gilbert, C. Feschotte, Genomic fossils calibrate the long-term evolution of hepadnaviruses. *PLOS Biology* **8** (2010).
24. M. B. Geuking *et al.*, Recombination of Retrotransposon and Exogenous RNA Virus Results in Nonretroviral cDNA Integration. *Science* **323**, 393-396 (2009).
25. T. Wicker *et al.*, A unified classification system for eukaryotic transposable elements. *Nature Reviews Genetics* **8**, 973-982 (2007).
26. M. Horie, K. Tomonaga, Paleovirology of bornaviruses: What can be learned from molecular fossils of bornaviruses. *Virus Research* **262**, 2-9 (2019).
27. M. Horie, Y. Kobayashi, Y. Suzuki, K. Tomonaga, Comprehensive analysis of endogenous bornavirus-like elements in eukaryote genomes. *Philosophical Transactions of the Royal Society B: Biological Sciences* **368**, 20120499 (2013).
28. V. A. Belyi, A. J. Levine, A. M. Skalka, Unexpected Inheritance: Multiple Integrations of Ancient Bornavirus and Ebolavirus/Marburgvirus Sequences in Vertebrate Genomes. *PLoS Pathogens* **6**, e1001030 (2010).
29. C. Gilbert *et al.*, Endogenous hepadnaviruses, bornaviruses and circoviruses in snakes. *Proceedings of the Royal Society B: Biological Sciences* **281**, 20141122

- (2014).
30. M. Dewannieux, C. Esnault, T. Heidmann, LINE-mediated retrotransposition of marked Alu sequences. *Nature Genetics* **35**, 41-48 (2003).
 31. C. Esnault, J. Maestre, T. Heidmann, Human LINE retrotransposons generate processed pseudogenes. *Nature Genetics* **24**, 363-367 (2000).
 32. A. J. Doucet, J. E. Wilusz, T. Miyoshi, Y. Liu, J. V. Moran, A 3' Poly(A) Tract Is Required for LINE-1 Retrotransposition. *Molecular Cell* **60**, 728-741 (2015).
 33. E. V. Koonin, Taming of the shrewd: novel eukaryotic genes from RNA viruses. *BMC Biology* **8**, 2 (2010).
 34. P. Klenerman, H. Hengartner, R. M. Zinkernagel, A non-retroviral RNA virus persists in DNA form. *Nature* **390**, 298-301 (1997).
 35. D. Rubbenstroth *et al.*, ICTV Virus Taxonomy Profile: Bornaviridae. *Journal of General Virology* **102** (2021).
 36. T. H. Hyndman, C. M. Shilton, M. D. Stenglein, J. F. X. Wellehan, Divergent bornaviruses from Australian carpet pythons with neurological disease date the origin of extant Bornaviridae prior to the end-Cretaceous extinction. *PLOS Pathogens* **14**, e1006881 (2018).
 37. M. Shi *et al.*, The evolutionary history of vertebrate RNA viruses. *Nature* **556**, 197-202 (2018).
 38. U. Schneider, Novel insights into the regulation of the viral polymerase complex of neurotropic Borna disease virus. *Virus Research* **111**, 148-160 (2005).
 39. J. C. De La Torre, Reverse-genetic approaches to the study of Borna disease virus. *Nature Reviews Microbiology* **4**, 777-783 (2006).
 40. B. Hoffmann *et al.*, A Variegated Squirrel Bornavirus Associated with Fatal Human Encephalitis. *New England Journal of Medicine* **373**, 154-162 (2015).
 41. K. Schlottau *et al.*, Variegated Squirrel Bornavirus 1 in Squirrels, Germany and the Netherlands. *Emerging Infectious Diseases* **23**, 477-481 (2017).
 42. D. Tappe *et al.*, Occupation-Associated Fatal Limbic Encephalitis Caused by Variegated Squirrel Bornavirus 1, Germany, 2013. *Emerging Infectious Diseases* **24**, 978-987 (2018).
 43. K. Korn *et al.*, Fatal Encephalitis Associated with Borna Disease Virus 1. *New England Journal of Medicine* **379**, 1375-1377 (2018).
 44. K. Schlottau *et al.*, Fatal Encephalitic Borna Disease Virus 1 in Solid-Organ Transplant Recipients. *New England Journal of Medicine* **379**, 1377-1379 (2018).
 45. R. Coras, K. Korn, S. Kuerten, H. B. Huttner, A. Ensser, Severe bornavirus-

- encephalitis presenting as Guillain–Barré-syndrome. *Acta Neuropathologica* **137**, 1017-1019 (2019).
46. C. Frank *et al.*, Emerging Microbes & Infections – Original Article: Human Borna disease virus 1 (BoDV-1) encephalitis cases in the north and east of Germany. *Emerging Microbes & Infections* 10.1080/22221751.2021.2007737, 1-19 (2021).
 47. S. L. Boatright-Horowitz, Avian Bornaviral Ganglioneuritis: Current Debates and Unanswered Questions. *Veterinary Medicine International* **2020**, 6563723 (2020).
 48. V. A. Costa *et al.*, Metagenomic sequencing reveals a lack of virus exchange between native and invasive freshwater fish across the Murray–Darling Basin, Australia. *Virus Evolution* **7** (2021).
 49. P. A. Schneider, A. Schneemann, W. I. Lipkin, RNA splicing in Borna disease virus, a nonsegmented, negative-strand RNA virus. *Journal of Virology* **68**, 5007-5012 (1994).
 50. B. Cubitt, C. Oldstone, J. Valcarcel, J. C. De La Torre, RNA splicing contributes to the generation of mature mRNAs of Borna disease virus, a non-segmented negative strand RNA virus. *Virus Research* **34**, 69-79 (1994).
 51. K. Tomonaga *et al.*, Identification of alternative splicing and negative splicing activity of a nonsegmented negative-strand RNA virus, Borna disease virus. *Proceedings of the National Academy of Sciences* **97**, 12788-12793 (2000).
 52. S. Kojima *et al.*, Splicing-Dependent Subcellular Targeting of Borna Disease Virus Nucleoprotein Isoforms. *Journal of Virology* **93**, e01621-01618 (2019).
 53. D. Rubbenstroth *et al.* (2018) One (1) new genus including one (1) new species in the family Bornaviridae (order Mononegavirales). (Tech. Rep. 2018.016M, ResearchGate).
 54. Zoonomia Consortium, A comparative genomics multitool for scientific discovery and conservation. *Nature* **587**, 240-245 (2020).
 55. S. Feng *et al.*, Dense sampling of bird diversity increases power of comparative genomics. *Nature* **587**, 252-257 (2020).
 56. Y. Kobayashi *et al.*, Exaptation of Bornavirus-Like Nucleoprotein Elements in Afrotherians. *PLOS Pathogens* **12**, e1005785 (2016).
 57. M. Hilbe *et al.*, Shrews as Reservoir Hosts of Borna Disease Virus. *Emerging Infectious Diseases* **12**, 675-677 (2006).
 58. M. Bourg *et al.*, Bicolored White-toothed Shrews as Reservoir for Borna Disease Virus, Bavaria, Germany. *Emerging Infectious Disease journal* **19**, 2064

- (2013).
59. M. D. Stenglein, E. B. Leavitt, M. A. Abramovitch, J. A. McGuire, J. L. DeRisi, Genome Sequence of a Bornavirus Recovered from an African Garter Snake (*Elapsoidea loveridgei*). *Genome Announc* **2** (2014).
 60. F. Pfaff, D. Rubbenstroth, Two novel bornaviruses identified in colubrid and viperid snakes. *Archives of Virology* **166**, 2611-2614 (2021).
 61. M. S. Springer, R. W. Meredith, J. E. Janecka, W. J. Murphy, The historical biogeography of Mammalia. *Philosophical Transactions of the Royal Society B: Biological Sciences* **366**, 2478-2502 (2011).
 62. H. Nishihara, S. Maruyama, N. Okada, Retroposon analysis and recent geological data suggest near-simultaneous divergence of the three superorders of mammals. *Proceedings of the National Academy of Sciences* **106**, 5235-5240 (2009).
 63. M. S. Springer *et al.*, Macroevolutionary Dynamics and Historical Biogeography of Primate Diversification Inferred from a Species Supermatrix. *PLoS ONE* **7**, e49521 (2012).
 64. J. I. Bloch, M. T. Silcox, D. M. Boyer, E. J. Sargis, New Paleocene skeletons and the relationship of plesiadapiforms to crown-clade primates. *Proceedings of the National Academy of Sciences* **104**, 1159-1164 (2007).
 65. E. Gheerbrant, A. Schmitt, L. Kocsis, Early African Fossils Elucidate the Origin of Embrithopod Mammals. *Current Biology* **28**, 2167-2173.e2162 (2018).
 66. M. A. Nilsson *et al.*, Tracking Marsupial Evolution Using Archaic Genomic Retroposon Insertions. *PLoS Biology* **8**, e1000436 (2010).
 67. M. D. B. Eldridge, R. M. D. Beck, D. A. Croft, K. J. Travouillon, B. J. Fox, An emerging consensus in the evolution, phylogeny, and systematics of marsupials and their fossil relatives (Metatheria). *Journal of Mammalogy* **100**, 802-837 (2019).
 68. C. Poux *et al.*, Asynchronous Colonization of Madagascar by the Four Endemic Clades of Primates, Tenrecs, Carnivores, and Rodents as Inferred from Nuclear Genes. *Systematic biology* **54**, 719-730 (2005).
 69. J. J. Jaeger, PALEONTOLOGY: Shaking the Earliest Branches of Anthropoid Primate Evolution. *Science* **310**, 244-245 (2005).
 70. M. Bond *et al.*, Eocene primates of South America and the African origins of New World monkeys. *Nature* **520**, 538-541 (2015).
 71. R. J. Gifford, Mapping the evolution of bornaviruses across geological timescales. *Proceedings of the National Academy of Sciences* **118**, e2108123118

- (2021).
72. E. C. Teeling, A Molecular Phylogeny for Bats Illuminates Biogeography and the Fossil Record. *Science* **307**, 580-584 (2005).
 73. J. Cui *et al.*, Low frequency of paleoviral infiltration across the avian phylogeny. *Genome Biology* **15** (2014).
 74. K. Ohshima *et al.*, Whole-genome screening indicates a possible burst of formation of processed pseudogenes and Alu repeats by particular L1 subfamilies in ancestral primates. *Genome Biology* **4**, R74 (2003).
 75. M. Horie, K. Tomonaga, Non-Retroviral Fossils in Vertebrate Genomes. *Viruses* **3**, 1836-1848 (2011).
 76. W. E. Johnson, Origins and evolutionary consequences of ancient endogenous retroviruses. *Nature Reviews Microbiology* **17**, 355-370 (2019).
 77. H. Kirsip, A. Abroi, Protein Structure-Guided Hidden Markov Models (HMMs) as A Powerful Method in the Detection of Ancestral Endogenous Viral Elements. *Viruses* **11**, 320 (2019).
 78. S. Kojima *et al.*, Virus-like insertions with sequence signatures similar to those of endogenous nonretroviral RNA viruses in the human genome. *Proceedings of the National Academy of Sciences* **118**, e2010758118 (2021).
 79. A. L. Greninger, A decade of RNA virus metagenomics is (not) enough. *Virus Research* **244**, 218-229 (2018).
 80. Y. Z. Zhang, Y. M. Chen, W. Wang, X. C. Qin, E. C. Holmes, Expanding the RNA Virosphere by Unbiased Metagenomics. *Annual Review of Virology* **6**, 119-139 (2019).
 81. N. A. O'Leary *et al.*, Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. *Nucleic Acids Research* **44**, D733-D745 (2016).
 82. C. Camacho *et al.*, BLAST+: architecture and applications. *BMC Bioinformatics* **10**, 421 (2009).
 83. K. Katoh, D. M. Standley, MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. *Molecular Biology and Evolution* **30**, 772-780 (2013).
 84. W. Shen, S. Le, Y. Li, F. Hu, SeqKit: A Cross-Platform and Ultrafast Toolkit for FASTA/Q File Manipulation. *PLOS ONE* **11**, e0163962 (2016).
 85. W. Bao, K. K. Kojima, O. Kohany, Repbase Update, a database of repetitive elements in eukaryotic genomes. *Mobile DNA* **6** (2015).
 86. A. A. Hagberg, D. A. Schult, P. J. Swart (2008) Exploring Network Structure,

- Dynamics, and Function using NetworkX. in *Proceedings of the 7th Python in Science Conference (SciPy2008)*, ed T. V. Gäel Varoquaux, and Jarrod Millman (Pasadena, CA USA), pp 11-15.
87. S. Kumar, G. Stecher, M. Suleski, S. B. Hedges, TimeTree: A Resource for Timelines, Timetrees, and Divergence Times. *Molecular Biology and Evolution* **34**, 1812-1819 (2017).
 88. M. J. Ankenbrand, S. Hohlfeld, T. Hackl, F. Förster, AliTV—interactive visualization of whole genome comparisons. *PeerJ Computer Science* **3**, e116 (2017).
 89. R. S. Harris (Improved pairwise alignment of genomic dna. (The Pennsylvania State University, 2007).
 90. J. E. Stajich, The Bioperl Toolkit: Perl Modules for the Life Sciences. *Genome Research* **12**, 1611-1618 (2002).
 91. A. S. Hinrichs *et al.*, The UCSC Genome Browser Database: update 2006. *Nucleic Acids Research* **34**, D590-D598 (2006).
 92. M. Haeussler *et al.*, The UCSC Genome Browser database: 2019 update. *Nucleic Acids Research* **47**, D853-D858 (2019).
 93. L.-T. Nguyen, H. A. Schmidt, A. Von Haeseler, B. Q. Minh, IQ-TREE: A Fast and Effective Stochastic Algorithm for Estimating Maximum-Likelihood Phylogenies. *Molecular Biology and Evolution* **32**, 268-274 (2015).
 94. S. Kalyaanamoorthy, B. Q. Minh, T. K. F. Wong, A. Von Haeseler, L. S. Jermiin, ModelFinder: fast model selection for accurate phylogenetic estimates. *Nature Methods* **14**, 587-589 (2017).
 95. D. T. Hoang, O. Chernomor, A. Von Haeseler, B. Q. Minh, L. S. Vinh, UFBoot2: Improving the Ultrafast Bootstrap Approximation. *Molecular Biology and Evolution* **35**, 518-522 (2018).
 96. G. Yu, D. K. Smith, H. Zhu, Y. Guan, T. T. Y. Lam, ggtree : an r package for visualization and annotation of phylogenetic trees with their covariates and other associated data. *Methods in Ecology and Evolution* **8**, 28-36 (2017).
 97. J. Huerta-Cepas, F. Serra, P. Bork, ETE 3: Reconstruction, Analysis, and Visualization of Phylogenomic Data. *Molecular Biology and Evolution* **33**, 1635-1638 (2016).

注釈

本学位論文は、以下の学術論文の内容に基づいて書かれたものである。

Junna Kawasaki, Shohei Kojima, Yahiro Mukai, Keizo Tomonaga, Masayuki Horie.

100-My history of bornavirus infections hidden in vertebrate genomes.

Proceedings of the National Academy of Sciences, Volume 118, Issue 20,
e2026235118, 2021.

DOI: 10.1073/pnas.2026235118

謝辞

本研究を遂行するにあたり、ご指導ご鞭撻いただいた朝長啓造教授、堀江真行教授（現・大阪府立大学）、並びにご協力いただいた皆様に心より感謝申し上げます。