

DNA修復・組換え遺伝子*recA*の進化的起源に関する  
遺伝学的及び分子系統学的解析

平成29年

明石 基洋

東京農業大学

博士論文

DNA修復・組換え遺伝子*recA*の進化的起源に関する  
遺伝学的及び分子系統学的解析

平成29年 修了

農学研究科バイオサイエンス専攻

明石 基洋

指導教授 千葉櫻 拓

# 目次

目次	p. 2
第1章 序論	p. 3
はじめに ～相同組換え酵素遺伝子はどこから来たのか～	
第2章 本論-1	p. 8
トランスポゾンIS256Bsu1を用いた実験の妥当性, およびトランスポゾンの転移と相同組換え酵素の関係性について.	
第3章 本論-2	p. 34
バクテリオファージMuにコードされる <i>muB</i> 遺伝子と相同組換え酵素遺伝子の分子系統学的関係性, およびトランスポゾンの転移における <i>muB</i> 遺伝子の相同組換え酵素遺伝子に対する相補性について.	
第4章 本論-3	p. 54
ゲノム上の突然変異, そのランダム性について.	
第5章 総合討論, および結論	p. 71
相同組換え酵素の進化プロセスについて.	
謝辞	p. 84
参考文献リスト	p. 85
付録	p. 99

# 第1章 序論

はじめに ～相同組換え酵素遺伝子はどこから来たのか～

# 第1章 序論

## 1. はじめに ～相同組換え酵素遺伝子はどこから来たのか～

紫外線 (UV-C) への曝露は、全ての生物にとってDNA障害の原因であり、生物が生存する上でいわゆる「自然選択」の要因となりうる環境条件の一つである。しかし、今日ではUV-Cの線量は地球を取り巻くオゾン層に吸収され、地表に届く量はほぼゼロとなっている。もしもこの防護壁がなければ、我々生き物は太陽の下で生きていくことは出来なかったであろう。化石資料に基づく調査によると、38億年前の地球には既に生命が存在したとされるが[1]、おそらくは紫外線に耐える能力は無く、光の届かない世界に潜んでいたと思われる。生物が紫外線の降りしきる中、より光の届く水面まで浮上し、更には地上へと進出するためには、光合成を介したオゾン層の形成という受動的なプロセスよりも前に、より積極的に紫外線に耐えるためのある能力の獲得が必須である。その能力とは、すなわち、DNA損傷修復能力である。

近藤宗平博士は、『分子放射線生物学』において、原子生命の誕生について、紫外線と窒素、酸素が反応することで合成された有機物を利用しつつ、紫外線を回避可能な深さの海底において原子生命は誕生、生活したというCarl Sagan (1934-1996) の仮説を紹介している[2,3]。ところが、本書において紫外線耐性、DNA修復機構の獲得について考察する記述は見られない[2]。筆者はこれを近藤博士から後進の者に宛てられた課題、問題提起であったと考える。そこで、本論文では、筆者がこれまで取り組んだ研究群をもとに、生物が紫外線耐性、DNA修復機構を獲得したプロセスの構築を試みる。

## 2. 相同組換え因子RecAとファージMuのDNA結合因子MuBの共通性

### 2-1. RecAについて

大腸菌 (*Escherichia coli*) のRecAは紫外線などの放射線によるDNA損傷修復において重要な役割を果たす。この酵素および遺伝子は、相同組換え活性を示す因子として非常に盛んに分子生物学的研究が進められてきた。SOS応答と呼ばれるDNA修復経路は、RecA以外にRecXやLexAなど複数の因子で構成されており、この経路は生物種を問わず広く保存されている[4-9]。RecAはATPと複合体を形成し、dsDNAが切断された際に生じた基質となるssDNAに結合し相補鎖を探索する相同組換え反応を担う[4,5]。通常、細胞内のssDNAはSSB(Single-strand binding proteins)により覆われているが、大腸菌において、RecAはRecFORによりSSBが除去されて露出したssDNA上にロードされる[10]。この時、RecOとRecRはRecAをロードする反応に必須だが、RecFは必須ではない[10]。枯草菌 (*Bacillus subtilis*) の場合、DNA損傷修復や相同組換えにはRecAが必要であるものの、プラスミドの形質転換にはRecAでなくRecOとRecUが必要である[11]。

*recA*は元々、水平伝播 (Horizontal gene transfer; HGT) の際にDNA相同組換えを引き起こす因子をコードする遺伝子として大腸菌で同定されたが[12,13]、その保存性は生物界にとどまらず、未知のウイルスあるいはバクテリオファージに由来する*recA*ホモログが存在することが知られている[14,15]。これらウイルスにコードされた相同組換え遺伝子に類似の配列は、その分子系統樹において各々クレードを形成し、RecAのクレードと根を共有している[14,15]。ここで、相同組換え遺伝子の系統関係に、ウイルスに保存された遺伝子が含まれることの意味について考える必要がある。ウイルスがこの遺伝子を系統的に保存しているということは、ウイルスの増殖過程に何らかの形で関与するということである。このウイルス由来の遺伝子群が既存の生物から獲得されたものならば、ウイルスのクレードは生物種由来配列のクレードに内包される。一方、ウイルス由来の配列群と生物種由来のクレードが分かれている場合、互いに独立した系統として進化したことを意味すると同時に、相同組換え関連遺伝子とウイルスの遺伝子に共通祖先配列が存在し、この共通祖先配列の宿主が、生物のみならずウイルスであったことを示唆する。後者の場合、相同組換え遺伝子の祖先配列が、ウイルスの感染 (水平伝播) によって生物にもたらされたことになる。一体、*recA*の祖先配列の由来は何であり、どのようにしてDNA修復能を獲得したのだろうか。

## 2-2. MuBとRecAの共通点, トランスポゾンとの関係性について

Muファージは, 感染した細胞のゲノム上で, コピー&ペースト型の複製を行うバクテリオファージであり, 1950年代後半にTaylorにより報告された[16]. このファージは, 約35kbのゲノム上に*muA*と呼ばれるDDE type Mutator like Transposaseをコードする遺伝子, *muB*と呼ばれるゲノム上におけるTransposaseのターゲット領域を決定するDNA結合タンパク質遺伝子を有している[17,18]. 以後の章で詳述するが, *recA*の祖先に近い配列が, Muファージにコードされている*muB*である可能性がある. というのも, *muB*と*recA*の遺伝子産物には, いくつもの共通する特徴が見られるのである. 例えば, NCBI Conserved Domain Databaseによると, MuBとRecAはともにP-loop dNTPase superfamilyに属する. また, いずれもDNA上でフィラメントを形成し, DNAから解離する際にATPを加水分解する[19].

トランスポゾンとはゲノム上を移動する塩基配列であり, 配列内に, 転移酵素(Transposase)をコードする. トランスポゾンゲノムの両端には逆向き反復配列が存在し(IR; Inverted repeat), 転移酵素はこの配列を認識して自身の塩基配列をゲノムから切り出す. 挿入配列(Insertion sequence; IS)はMuファージなどとは異なり, 転移酵素と挿入配列のみで構成された最小単位のトランスポゾンである[20]. 挿入配列(IS)にコードされる転移酵素は, 活性中心となる3つのアミノ酸のモチーフを保存しており(DDEモチーフ), このモチーフのアミノ酸配列上の位置が挿入配列の種類(IS family)によって異なる[20].

納豆菌(*Bacillus subtilis natto*)には, IS256*Bsu1*と呼ばれるトランスポゾンが存在する[21]. この挿入配列(IS)はIS256 familyに属し, 先に述べたMuAと同様, DDE type Mutator like Transposaseをコードしている[22-25]. 納豆菌のゲノム上には8~10コピーのIS256*Bsu1*が確認されており, しかもIR配列の両側に重複配列(後述)が確認されているため, IS256*Bsu1*は転移活性のあるISであると考えられる[21]. これまで, 大河内らにより本ISの転移を検出する系(Jumping cat assay)が枯草菌168株(*Bacillus subtilis* 168)で作製され, 枯草菌内でこのISの転移に関与する遺伝子(Host factor)の探索が行われてきた[26]. 筆者は, 枯草菌にIS256*Bsu1*の細胞内における転移反応に*recA*が必須であるという事実(茂木ら[27])から研究を開始, その必須性の要因を遺伝学的に検証した. その過程で, Muバクテリオファージにコードされる遺伝子*muB*が, IS256*Bsu1*の転移における*recA*の必須性を相補する事実を発見した[28]. そこで, 細胞性生物の相同組換え酵素遺伝子と*muB*の分子系統学

的關係性を調査し，細胞性生物の相同組換え酵素遺伝子獲得にバクテリオファージの感染が関与している可能性があることを明らかとした[28].

### 3. 本論文の狙い

本論文では，以上に概説した研究結果を中心に，相同組換え酵素遺伝子を生物が獲得した過程を明らかにすることを目指す. 本章に続く第2章では，IS256*Bsu1*にコードされる転移酵素（トランスポゼース）による転移の検出方法について述べる. また，相同組換え酵素RecAの如何なる機能が転移反応に重要であると考えられるか，分子遺伝学的実験により検証した結果を提示，考察する. 第3章ではまず，本研究においてバクテリオファージMuにコードされる*muB*に着目した経緯について述べる. 続いて，*muB*が生物一般に見られる相同組換え酵素遺伝子との分子系統学的関係性について解析した結果を示し，そこから見出された変異の特徴についても言及する. 最後に，*muB*のIS256*Bsu1*の転移及びUV耐性における*recA*への相補性を確認した結果を示し，*recA*または*muB*が，IS256*Bsu1*の転移においてどのような働きをしているのかについて考察する. 第4章では第3章における相同組換え遺伝子の系統解析結果を踏まえた上で，より発展的な話題として，筆者が行なった生物ゲノムのGC含量と高頻度突然変異株（通称：ミューテーター）の関係に関する研究，及び，突然変異のランダム性に関わるシミュレーションを紹介し，ゲノム上の突然変異の実際について考察する. 最後に，第5章では，これらの実験結果に基づき，生物が相同組換え酵素遺伝子を獲得した過程を推定し，本研究の締めくくりとする.



# 第2章 本論-1

トランスポゾンIS256*Bsu*1を用いた実験の妥当性、  
およびトランスポゾンの転移と相同組換え酵素の関係性について。

# IS256Bsu1の転移頻度と培地種の関係について

本研究に登場する培地，菌株および薬剤の使用条件などについて紹介する．また，IS256Bsu1の転移頻度（Transposition Frequency; TPF）の検出系であるJumping cat assayの構築方法および検出原理について概説する．続いて，IS256Bsu1の転移と培地条件の関係性について述べる．

## <材料, 方法>

### 1-1. 菌株, プラスミド, 培地条件について

本研究に用いた枯草菌は，枯草菌168株 (*Bacillus subtilis* Marburg 168) をベースとしている．この枯草菌は大腸菌と並んで古くから実験に用いられてきた菌であり，様々な分子生物学的な知見が蓄積されている．この株が歴史的に実験に多用された要因として，毒性がないことや形質転換が容易である点が考えられるが，トレンドの問題でもあるため，詳細は定かではない．近年では168株のみならず，その近縁の枯草菌のゲノム情報も多数公開されており，より多角的な研究が可能となった．これらに加え，本研究において168株が使用されるに至った理由は，この株がトランスポゾンフリーで実験が可能なバクテリアであったためである．枯草菌を培養した培地はLuria-Bertani (LB) 培地，枯草菌の形質転換に使用する形質転換誘導 (competence induction ; CI) 培地[29]，枯草菌の孢子形成を誘導する2xSG培地[30]の3種類で，培養温度は37°Cである．後述するJumping cat assayの構築にあたりプラスミドの増幅のために，大腸菌K-12株に由来するDH10B株 (*Escherichia coli* DH10B [F<sup>-</sup>, *mcrA*,  $\Delta$  (*mrr-hsdRMS-mcrBC*) ,  $\Phi$ 80*dlacZ*,  $\Delta$ M15,  $\Delta$ *lacX74*, *deoR*, *recA1*, *araD139*,  $\Delta$ (*ara leu*)7697, *galU*, *galK*,  $\lambda$ -, *rpsL*, *endA1*,*nupG*]) を使用している[31]．大腸菌の培養はLB培地を使用し，培養温度は37°Cとした．納豆菌 (*Bacillus subtilis* BEST195 [32]) のゲノムDNAは，IS256Bsu1の鋳型として使用した．大腸菌-枯草菌のシャトルベクタ pDR111aは，

Jumping cat assayシステムの構築に利用されている[33]. なお, 本プラスミドの選択条件は, 大腸菌ではampicillin, 枯草菌ではspectinomycinである. 実験で使用した全ての抗生物質の濃度は次の通りである (chloramphenicol (Cm) , 5  $\mu\text{g}/\text{mL}$ ; ampicillin, 100  $\mu\text{g}/\text{mL}$ ; spectinomycin, 50  $\mu\text{g}/\text{mL}$ ; erythromycin, 1  $\mu\text{g}/\text{mL}$ ; tetracycline, 15  $\mu\text{g}/\text{mL}$ ) .

## 1-2. Jumping cat assay

まず, トランスポゾンの*in vivo*検出系であるJumping cat assayについて概説する. Jumping cat assayは 大河内 悠貴 氏 (平成22年度修士課程修了) および茂木俊丞氏 (平成23年度修士課程修了) らにより構築された検出系である[26,27]. 本実験系では転移するのはIS256*Bsu1*ではなく, 枯草菌のShine-Dalgarno (SD) 配列を接続したchloramphenicol耐性遺伝子 (*cat*と表記, chloramphenicol-acetyltransferaseをコードする) を, 40 bpからなるIS256*Bsu1*の逆向き反復配列 (Inverted repeat; IRR / IRL) で挟んだ配列である (mini-ISと呼ぶ) .ただし, この*cat*遺伝子はプロモータ配列を持たない. 一方, IS256*Bsu1*にコードされるトランスポゼース遺伝子 (*tnp*) はラクトースのアナログであるIPTG (Isopropyl  $\beta$ -D-1-thiogalactopyranoside) 誘導性のプロモータ配列P<sub>hyper-spank</sub>に繋がれている[34]. これらの構造は枯草菌ゲノム上の*amyE*遺伝子内に挿入されている. *tnp*遺伝子が発現し, mini-ISが転移してゲノム上の他の遺伝子内に挿入されると, 転写融合が起きた場合のみ細胞の表現型がchloramphenicol耐性 (Cm<sup>r</sup>) となり, コロニーとして検出される. 言い換えれば, 転写融合が起きなければ, mini-ISが転移しても検出されない (図1) . なお, Jumping cat assayの名前は, *cat*遺伝子がmini-ISとしてゲノム上をジャンプ (転移) する事に由来する.

続いて, Jumping cat assayシステムの構築方法について述べる. 以下に記載するプライマーの配列は付録5に記載した. まず, mini-ISの導入する構造内での転写/翻訳を防止するため, ターミネーター配列の構築を行った. 配列は, 枯草菌の転写制御配列のデータベースであるDBTBS[35]を参照し, *glyQ/StRNA*アテニューエータ配列と*xkd*オペロンのターミネーター配列を選択した. これらの配列は, 枯草菌168株のゲノムを鋳型に, それぞれ適切に設計したプライマー (A1-FとA1-R, A2-FとA2-R) を用いてPCRにより増幅した. 増幅した二つのDNA断片は, A1-FプライマーとA2-Rプライマーを用いて, Recombinant PCR法[36]により結

合した。得られたPCR断片をNheI/SphIにより制限酵素処理し、同じ酵素で処理したpDR111aと結合し、pDR111a-T2を得た。尚、プラスミドpDR111aは枯草菌の*amyE*領域に挿入可能な組込型ベクターである。

続いて、IS転移を検出するmini-IS遺伝子カセットを構築した。mini-IS配列はクロラムフェニコール耐性遺伝子（*cat*遺伝子）の両側に、IS256*Bsu1*の両端に存在する逆向き反復配列（Inverted repeat sequence）が繋がれた構造をしている。*cat*遺伝子はpC194 [37]に由来する配列と、枯草菌のShine-Dalgarno (SD)配列から成るが、プロモーター配列は含まれない。まず、*cat*遺伝子をpBEST4C [38]からB1-FプライマーとB1-Rプライマーを用いてPCRにより増幅した。得られたDNA断片を、IS256*Bsu1*の逆向き反復配列（IRR及びIRL）とその末端にNheI/SalI認識配列を付加したプライマー（B2-F/B2-R）を用いてPCRにより増幅した。一方で、IS256*Bsu1*のトランスポゼース遺伝子（*tnp*）は枯草菌BEST195株（納豆菌）のゲノムを鋳型にC-FプライマーとC-Rプライマーを用いて増幅した。尚、*tnp*遺伝子の第一コドンはGTGであるが、Jumping cat assayが確実に機能するようATGに変更している。

得られたmini-IS断片と*tnp*遺伝子断片は、それぞれNheI/SalI及びHindIII/SalIにより制限酵素処理し、pDR111a-T2のNheI/HindIII 内に組み込んだ。最終産物として得られた、ターミネーター配列、mini-IS配列、*tnp*遺伝子を持つプラスミドを用いて、枯草菌168株を形質転換した。*amyE*領域への挿入は、スターチプレートを用いた菌株の澱粉分解活性（アミラーゼ活性）の欠損により確認した。最終的に作出されたJumping cat assay用変異株は、全部で18株である（付録を参照）。株は全て凍結保存した。

Jumping cat assayの実験手順は次の通りである。凍結した株をLB寒天培地に塗り広げ、16時間ほど25°Cで前培養した後、吸光度（600 nm）がおおよそ0.05となるよう5 mlのLB、CIあるいは2xSG培地を入れたL字管（底の長さが7 cmのもの）に植え、37°Cで24時間振盪培養（48 rpm）した（いずれの培地条件でも、IPTGを1 mMの濃度で加えている）。この培養液を10<sup>5</sup>または10<sup>6</sup>に希釈してLB寒天培地に、希釈していない培養液を5 µg/mLのchloramphenicolを含むLB寒天培地に各々3枚ずつ塗抹した。ただし、2xSG培地で培養したサンプルは、胞子を形成し、かつ表現型がCm<sup>r</sup>であるものの頻度を調べるため、80°Cで10分培養液を加熱した後に培地上に塗抹した。培養液を塗抹したプレートは37°Cで24時間培養

し、コロニー数を数えた後、1 mlあたりのColony forming unit (CFU) の平均値を求めた。mini-ISの転移頻度 (transposition frequency; TPF) は、 $TPF = [5 \mu\text{g/mL Cmを含むLBプレートから算出されたCFU}] / [\text{Cmを含まないLBプレートから算出されたCFU}]$ として定義した。各種菌株におけるTPFの計測は、それぞれ3回以上行った。また、TPFの比較は、Jumping cat assay用の構造のみを導入した*B. subtilis* (wild type; WTと表記する) との間で行った。WTとmutantのTPFの統計的な有意差の有無は、並べ替えBrunner-Munzel検定を用いて確認した[39]。

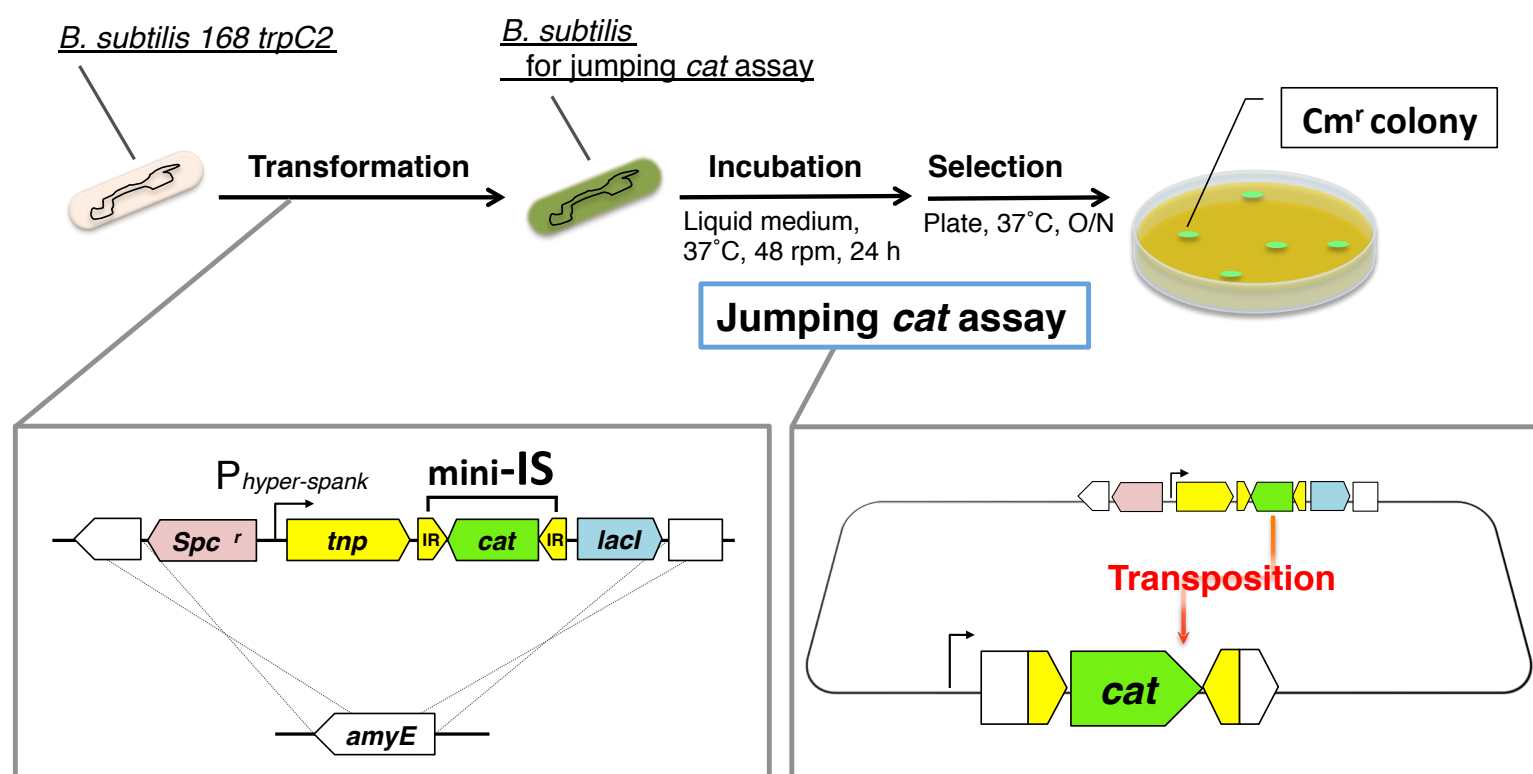


図1. Jumping cat assay 株の構築及び転移頻度 (TPF) 測定までの概略図

上段左側の細胞は枯草菌168株 (*B. subtilis* 168 *trpC2*) , 上段中央の細胞はJumping cat assay systemが形質転換された枯草菌168株 (*B. subtilis* for jumping cat assay) である. この細胞を培養, chloramphenicolを含む培地で選択すると, mini-ISの転移した細胞に由来するコロニー (上段右側, Cm<sup>r</sup> colony) が得られる. 下段左側枠内の図は, Jumping cat assay systemが*amyE*に挿入される様子を示す. *Spc<sup>r</sup>*はSpectinomycin耐性遺伝子であり, 枯草菌に Jumping cat assay systemを形質転換する際の選択マーカである. *P<sub>hyper-spank</sub>*はIPTG誘導性のプロモータであり, *tnp*遺伝子 (IS256*Bsu1*のトランスポゼース遺伝子) の5'末端側と接続している. mini-ISは*cat*遺伝子 (chloramphenicol耐性遺伝子) とその両端に繋がれた IS256*Bsu1*の逆向き反復配列 (IR) から成る. *lacI*遺伝子はIPTG及び*P<sub>hyper-spank</sub>*へと結合するLacリプレッサーをコードする. 下段右側枠内の図は, Jumping cat assay system内のmini-ISが, 枯草菌ゲノム上のある遺伝子内に挿入され, かつ検出可能な場合の例を示す. mini-ISの転移は, 挿入された遺伝子との転写融合により検出可能であるため, mini-IS内の*cat*遺伝子と挿入された遺伝子の転写される向きが一致している必要がある.

## <結果, 考察>

### 1-3. mini-IS転移頻度の培地間比較

まず, WTのmini-ISの転移頻度を3種類の異なる培地 (LB, CI, 2xSG) を用いて比較, CI培地における測定値との有意差検定を実施した. 転移頻度の測定方法は先の項をを参照する. ただし, LB培地での転移頻度は茂木の値を参照している[27]. LB培地は細菌を用いた実験で一般的に用いられる, 栄養を豊富に含む培地である. 一方, CI培地や2xSG培地はLB培地と比較して栄養が少ない. それぞれ異なる目的でデザインされたもので, CI培地は枯草菌の形質転換に使用する培地[29], 一方で2xSG培地は枯草菌に孢子を形成させるための培地である[30]. 実験の結果, CI培地で培養したサンプルが最も高い転移頻度を示した (図2). 更に, 2xSG培地で孢子を形成し, かつCm<sup>r</sup>となった細胞の割合は, 2xSG培地で観察された全Cm<sup>r</sup>菌の42.2%であった. 2xSG培地において孢子を形成した細胞の割合は87%であった. 即ち, mini-ISが転移したと考えられる細胞の半数が, 孢子を形成していない事が分かった. 以上を踏まえると, mini-ISの転移は栄養源が少なく, かつ孢子を形成していない細胞において頻繁に生じていると考えられた. 従って, mini-ISの転移を観察するためには転移頻度が最も高いCI培地を用いる事が適当であると考え, 以後のJumping cat assayではCI培地を用いる事にした.

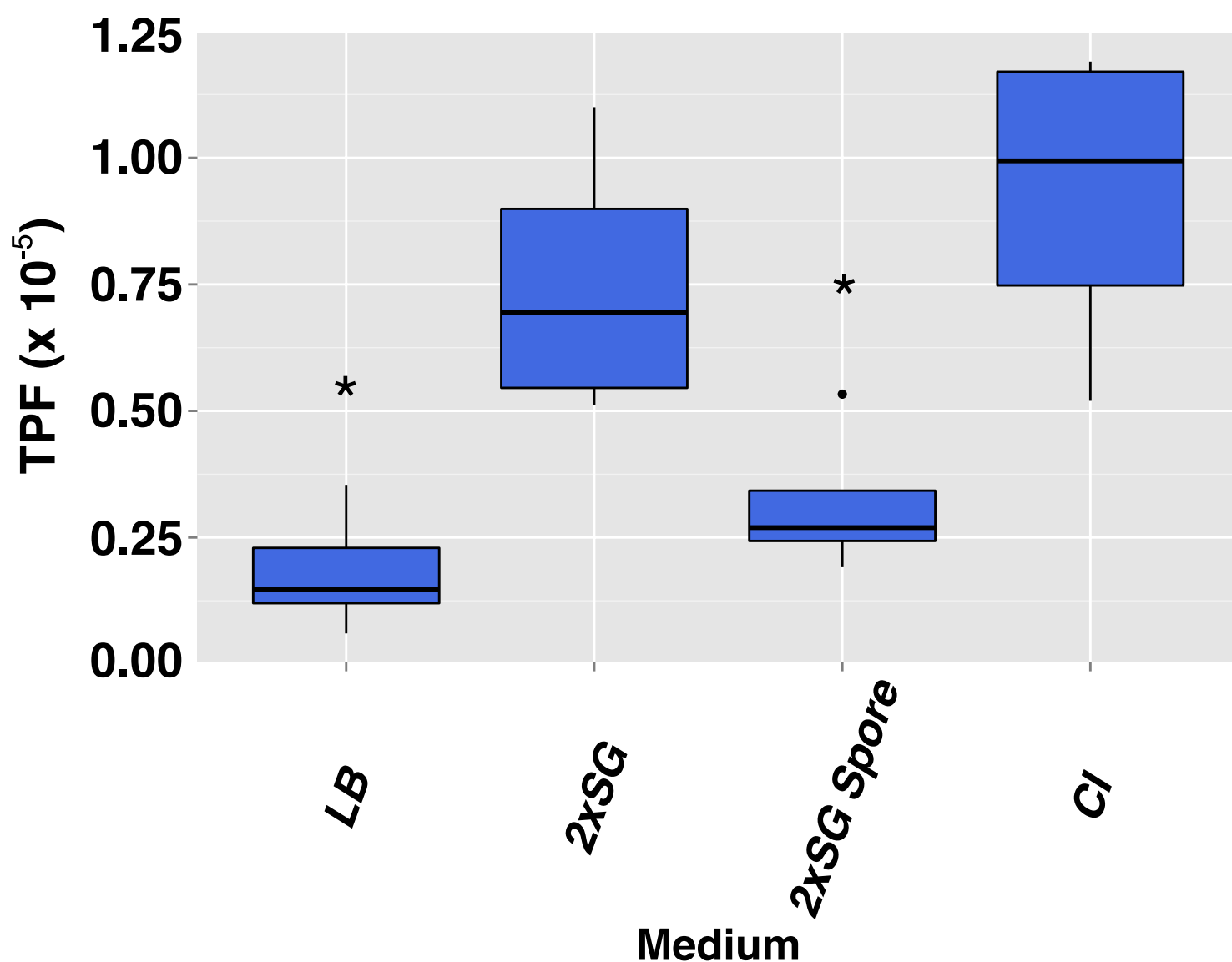


図2. mini-IS転移頻度の培地間比較

Jumping cat assay systemを導入した枯草菌を用いた, mini-IS転移頻度の培地間比較結果をBox plotで示す. 箱の上辺は第一四分位点, 中央の太線は中央値, 下辺は第三四分位点を示す. 箱に対して垂直な直線の上端は最大値, 下端は最小値, ドットは外れ値を意味する. TPF (縦軸) は転移頻度, Medium (横軸) にあるLB, 2xSG, CIはそれぞれ培地の名称である. 2xSG Sporeは, 孢子形成培地で孢子を形成した細胞における転移頻度である. アスタリスク (\*) は, 並べ替えBrunner-Munzel検定により, CI培地を用いた場合の転移頻度と比較して有意差が認められたものを示す ( $p < 0.05$ ). 尚, LBの転移頻度は茂木 俊丞氏の測定値を参照した[27].



# IS256*Bsu*1の重複配列の検出

ここでは、Jumping cat assayシステムを有する枯草菌に由来するCm<sup>r</sup>コロニーが、本当にmini-ISの転移に由来するのかを検証した研究について記載する。

## 2-1. ISと重複配列について

IS256 familyに属す挿入配列は、他の多くの挿入配列と同様[21,40], ISの両側に短い重複配列 (Direct repeat sequence; DR) を生じる[20]. もし、Jumping cat assayにより得られたCm<sup>r</sup>コロニーがmini-ISの転移に由来するならば、これらのコロニーにおけるmini-ISの挿入位置の両側にも重複配列が検出されるはずである。一方で、mini-ISが相同組換えなどにより他の遺伝子内に転座あるいは挿入された結果Cm<sup>r</sup>コロニーが生じたならば、重複配列は生じず、検出されるmini-ISの近傍に、*amyE*内のmini-IS周辺の配列が見出されると考えられる。つまり、重複配列の有無をもって、Jumping cat assayの転移頻度検出系としての有効性は担保される。

<材料, 方法>

## 2-2. 重複配列の確認方法

Jumping cat assayの有効性を確認するため、mini-ISの転移の際に生じる重複配列の検出を行なった。まず、LB培地及びCI培地で培養したJumping cat assayシステム導入枯草菌より、クロラムフェニコール耐性菌を6個ずつ、計12個を単離し、DNAを抽出した。まず、重複配列の片側を検出するため、Takara LA-PCR<sup>TM</sup> *in vitro* Cloning Kit (Takara Bio Inc., Otsu, Japan)に付属のHindIIIカセットを、HindIII処理したゲノムDNAに付加した。以下で述べるプライマーの配列は付録5を参照のこと。最初のPCRは、HindIIIカセットに特異的なC1プライマーと*cat*遺伝子に特異的なS1プライマーまたはS3プライマーを使用した。このPCR反応液

を鋳型として、HindIIIカセットに特異的なC2プライマーと*cat*遺伝子に特異的なS2プライマーまたはS4プライマーを使用してPCRを行なった。C1プライマーとC2プライマーはTakara LA-PCR™ *in vitro* Cloning Kitに付属のものである。また、S1からS4の4種類のプライマーについては、S1とS2が*cat*遺伝子に対して逆向き、S3とS4が順向きに設計されている。この検出法はNested PCR法であり2回目のPCRは1回目のPCRの内側領域となっている。続いて、特定されたmini-IS挿入領域のゲノム情報に基づき、次のプライマーによりmini-ISの両側にある重複配列を検出した(*yabC*: A-F/A-R, *queF*: B-F/B-R, *tasA*: C-F/C-R, *bceB*: D-F/D-R, *comP*: E-F/E-R, *atpH*: F-F/F-R)。尚、シーケンス解析はMACROGEN japan (<http://www.macrogen-japan.co.jp>)に委託した。

## <結果, 考察>

### 2-3. Cm<sup>r</sup>コロニーのmini-ISにおいて検出された重複配列について

以上の実験により、最終的に6種類の重複配列が検出された。解析に用いた12個のコロニーに由来するDNAの内、6コロニーからmini-ISの転移が検出され、そのうち5コロニーからは一種類の転移が検出され、1つのコロニーからは二種類の転移が検出された。検出された重複配列のうち、2つが同一であった。いずれの重複配列の長さも8 bpであったが、その塩基配列に保存性は見られなかった(表1)。この重複配列の長さは、黄色ブドウ球菌(*Staphylococcus*)の持つIS256のものと同様であった[41]。また、mini-ISの挿入された枯草菌のゲノム上の領域に規則性は無いようであるが(付録1を参照)、全てのmini-ISにおける*cat*遺伝子のORFは、挿入された遺伝子のORFの向きと一致していた(表1)。従って、Jumping cat assayにより生じるCm<sup>r</sup>コロニーは、mini-ISが転移し、転写融合により*cat*遺伝子が発現した結果得られたものである事が分かった。つまり、Jumping cat assayは、IS256*Bsu1*の転移頻度検出系として機能する事が示された。

Direct repeat (5'→3')	Position	Direction	Inserted gene	Product
<b>TTTGAAGC</b>	44859 - 44866	+	<i>yabC</i>	<i>16S rRNA methyltransferase</i>
<b>TGTTTATT</b>	1441408 - 1441415	-	<i>queF</i>	<i>nitrile reductase</i>
<b>ACTCCTCT</b>	2553019 - 2553026	+	<i>tasA</i>	<i>major component of biofilm matrix</i>
<b>GCGGAATG</b>	3109582 - 3109589	+	<i>bceB</i>	<i>bacitracin ABC transporter (permease)</i>
<b>GAGATGAA</b>	3255545 - 3255552	+	<i>comP</i>	<i>two-component sensor kinase</i>
<b>AAATGCAT</b>	3785764 - 3785771	+	<i>atpH</i>	<i>ATP synthase (subunit delta)</i>

**表1. mini-ISにおいて検出された重複配列**

mini-ISの転移により生じた重複配列 (Direct repeat) 及び, 枯草菌ゲノム (GenBank: CP010052.1) における挿入位置 (Position), 挿入された遺伝子 (Inserted gene) とその遺伝子産物 (Product) を示す. Directionは, 解読された枯草菌ゲノム (CP010052.1) の座標に対する挿入されたmini-IS内の*cat*遺伝子の向きを示す.

# IS256*Bsu1*の転移頻度とコンピテンス関連遺伝子/相同組換え遺伝子の関係性

ここでは、IS256*Bsu1*の転移に関与する遺伝子の探索を行った結果について述べる。対象となる遺伝子の探索は、コンピテンス関連遺伝子と相同組換え遺伝子に焦点を絞って行われた。

## 3-1. コンピテンス関連遺伝子とmini-ISの転移頻度の関係性

CI培地を用いてJumping cat assayを行った場合、他の培地と比べてmini-ISの転移が高頻度に検出された（図2）。CI培地は形質転換を誘導する培地（competence-induction medium）である事を考慮すると、mini-ISの転移に関与する枯草菌の遺伝子は、コンピテンスに関与するものである可能性が高いと考えられる。コンピテンスとは、自然形質転換、つまり細胞外に存在するDNAを自ら細胞内に取り込み、形質転換を引き起こす一連の反応のことを指す。枯草菌では、菌の密度が上昇すると、ComXと呼ばれるペプチド（フェロモン）が放出されるが、このシグナルを、ComP, ComA二つの因子のリン酸基転移反応に基づくシグナル伝達経路を介して、ComSの発現が促進される。この因子はコンピテンス関連遺伝子の転写因子ComKの分解を抑制する因子である。このComSの発現に加え、栄養飢餓によりComKが発現し、DNAの取り込みに関与する遺伝子の転写が起こる[42]。ComKにより転写される遺伝子は多数存在するが、その内で最も重要な遺伝子は相同組換え因子をコードする*recA*である。また、Rokは*comK*の転写を抑制する因子であるが、ComKによって*rok*の転写は阻害されている。

## <材料, 方法>

まず, 形質転換に関与する遺伝子五つ (*comP*, *comA*, *comK*, *rok*, *recA*) をそれぞれ別途に破壊した枯草菌 (Jumping cat assayシステムを搭載したもの) を作製した. 作製された株の構造は付録5を参照のこと. *comP*, *comA*, *comK*, *recA*, *rok*破壊株は, 該当する遺伝型を持つ枯草菌のゲノムをJumping cat assay systemを含む枯草菌168株に形質転換し (NBS290: *trpC2 recA::tet*, NBS295: *trpC2 comK::tet*, NBS798: *trpC2 comP::tet*, NBS799: *trpC2 comA::tet*, NBS1193: *trpC2 rok::tet*), 薬剤入りの培地で選択, 単離し, 目的の株を得た. 続いて, Jumping cat assayにより, CI培地における野生株 (WT) とこれらの株の転移頻度 (TPF) を比較, 有意差検定を実施した. 転移頻度の測定方法は第2章第1節を参照する. 尚, コンピテンス関連遺伝子の測定値は茂木俊丞氏の値を参照している[27].

## <結果, 考察>

結果, コンピテンスの誘導に関与する遺伝子 (*comP*, *comA*, *comK*, *rok*) の破壊株では, WTと比較してTPFの顕著な低下が見られた ( $p < 0.01$ , 図3). さらに, 形質転換の際にDNA相同組換えを担う因子をコードする*recA*の破壊株のTPFは, 野生株と比べて約0.001倍まで低下した ( $p < 0.01$ , 図3). つまり, *recA*を破壊すると, mini-ISの転移が起こらないと言える. 枯草菌では, *recA*の転写は主にComKに依存しているため[43], 形質転換能の誘導に関与する遺伝子の破壊株においてWTよりもTPFが低下したのは, これらの遺伝子の欠損による正常なコンピテンス誘導の阻害により, *recA*の発現量が低下したためであると解釈できる.

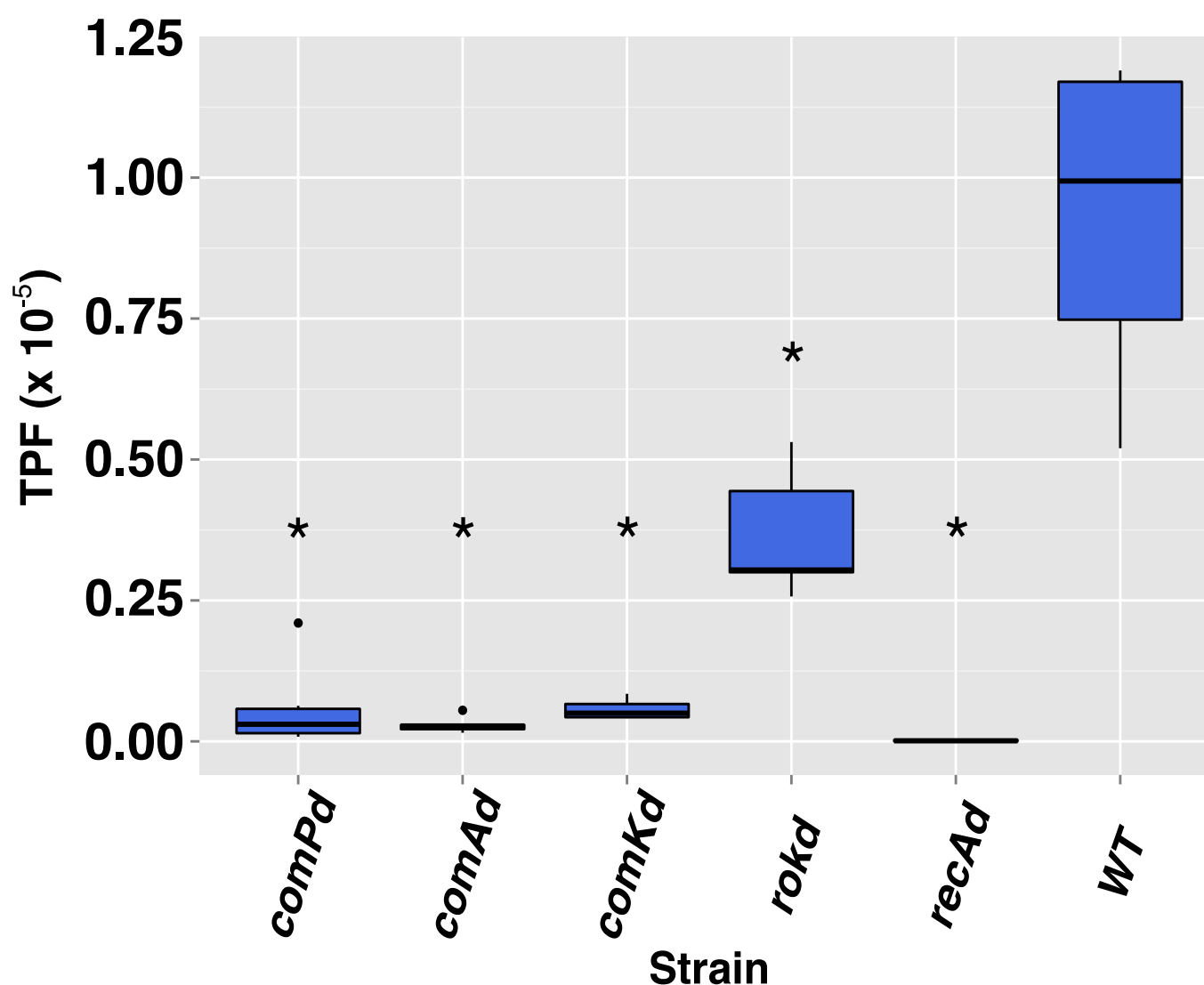


図3. コンピテンス関連遺伝子とmini-ISの転移頻度の関係性

mini-IS転移頻度をコンピテンス関連遺伝子欠損株を用いて測定した結果をBox plotで示す。箱の上辺は第一四分位点，中央の太線は中央値，下辺は第三四分位点を示す。箱に対して垂直な直線の上端は最大値，下端は最小値，ドットは外れ値を意味する。TPF（縦軸）は転移頻度，Strain（横軸）はそれぞれ欠損した遺伝子の名称を示す（*comPd*: *comP*欠損，*comAd*: *comA*欠損，*comKd*: *comK*欠損，*rokD*: *rok*欠損，*recAd*: *recA*欠損，WT: 野生型）。アスタリスク（\*）は，並べ替えBrunner-Munzel検定により，WTの転移頻度と比較して有意差が認められたものを示す（ $p < 0.05$ ）。尚，*comKd*, *comPd*, *comAd*, *rokD*, *recAd*の転移頻度は茂木 俊丞氏の測定値である[27]。

### 3-2. 相同組換え遺伝子とmini-ISの転移頻度の関係性

続いて、相同組換え反応とmini-ISの転移について検証するため、*recA*と共に相同組換えに関与する因子をコードする*recO*と*recU*に注目した。RecOは相同組換え反応の際、RecAをssDNA上にロードするのを補助する因子の一つである[10]。また、RecUはRecOと共に、プラスミドの形質転換に関与する[11]。相同組換え反応がmini-ISの転移に関与する場合、これらの因子の破壊でもmini-ISの転移が検出されなくなるか、あるいは転移頻度の著しい低下が見られるはずである。

#### <材料, 方法>

そこで、実際にコンピテンス関連遺伝子での実験と全く同様の方法で*recO*や*recU*を破壊した株を作製した。作製方法は次の通りである。以下のプライマーの配列及び作製された株の構造は付録を参照のこと。まず、枯草菌168株のゲノムを鋳型に、F1-F, F1-R及びF3-F, F3-Rのプライマを用いて、また、NBS909 (*trpC2 odhB::erm*) のゲノムを鋳型にF2-F, F2-Rのプライマを用いて、PCR反応によりDNA断片を作製した。続いて、得られた三種類のDNA断片を、F1-F, F3-Rのプライマを用いて、リコンビナントPCR法[36]により一つの断片に繋いだものを作製した。得られたDNA断片を、Jumping cat assay systemを含む枯草菌168株に形質転換し、薬剤入りの選択培地で選択、単離し、目的の株を得た。続いて、Jumping cat assayにより、作製した株の転移頻度を測定した。転移頻度の測定方法は第2章第1節を参照する。

#### <結果, 考察>

*recO*や*recU*の破壊株における転移頻度は、いずれもWTより高い値を示した ( $p < 0.01$ , 図4)。*recA*がmini-ISの転移に必須であるにも関わらず、*recO*や*recU*の破壊でTPFがWTより高くなることから、当初の予想に反して、これらの遺伝子産物はmini-ISの転移を何らかの形で阻害している可能性が考えられた。

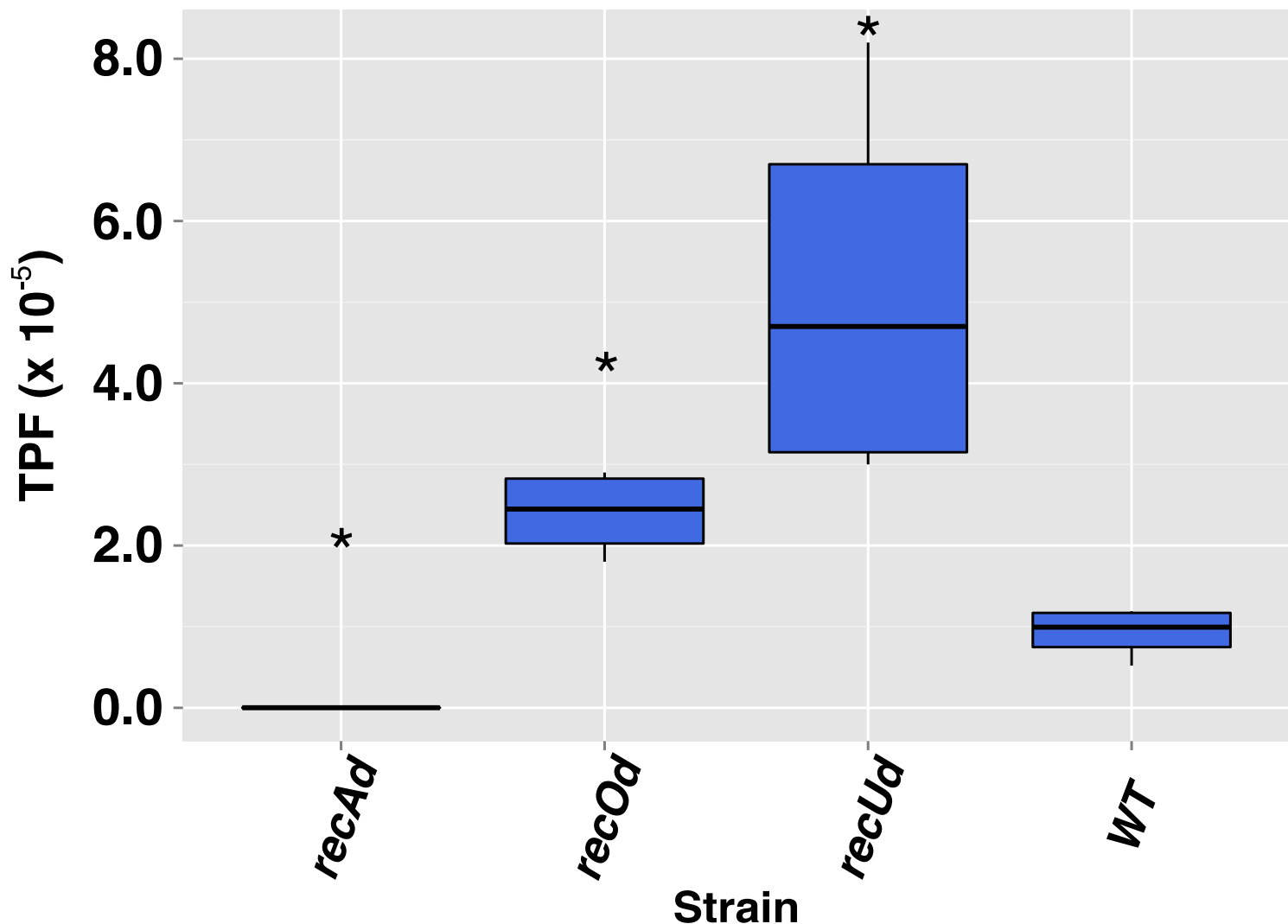


図4. 相同組換え関連遺伝子とmini-ISの転移頻度の関係性

mini-IS転移頻度を相同組換え関連遺伝子欠損株を用いて測定した結果をBox plotで示す。箱の上辺は第一四分位点，中央の太線は中央値，下辺は第三四分位点を示す。箱に対して垂直な直線の上端は最大値，下端は最小値を意味する。TPF（縦軸）は転移頻度，Strain（横軸）はそれぞれ欠損した遺伝子の名称を示す（*recAd*: *recA*欠損，*recOd*: *recO*欠損，*recUd*: *recU*欠損，*WT*: 野生型）。アスタリスク（\*）は，並べ替えBrunner-Munzel検定により，野生株（*WT*）の転移頻度と比較して有意差が認められたものを示す（ $p < 0.05$ ）。尚，参考のため，茂木 俊丞氏の*recAd*の測定値を並べて図示した[27]。



### 3-3. 小括: *recA*はmini-ISの転移に必須だが, 相同組換え反応は必須でない.

改めて, これまでの実験結果を整理する. mini-ISの転移頻度は栄養の少ない条件, 中でもコンピテンスを誘導した際に最も高い値を示した. また, mini-ISの転移には重複配列が検出されており, 相同, あるいは非常によく似た塩基配列の相同組換えによってmini-IS上の*cat*遺伝子が発現している訳ではないことは明らかである. さらに, コンピテンスを誘導する遺伝子を破壊した場合に転移頻度の低下がおこることが確認されたが, その原因は結局のところ*recA*が正常に発現しなくなったためであろうと推察される. 何故ならば, *recA*を欠損すると, mini-ISの転移は起こらない. つまり, *recA*がmini-ISの転移には必須なのである. 一方, *recA*の転移への関与の仕方は, 驚くべきことに, 相同組換え反応ではない可能性が高い. 何故ならば, 相同組換え反応をRecAと共に進める遺伝子の破壊株 (*recO*や*recU*) では, むしろ野生株と比較して転移頻度が高いためである. 以上から, *recA*がmini-ISの転移に, 相同組換え反応とは異なる方法で関与する事が示唆された. この発見は意外であった.

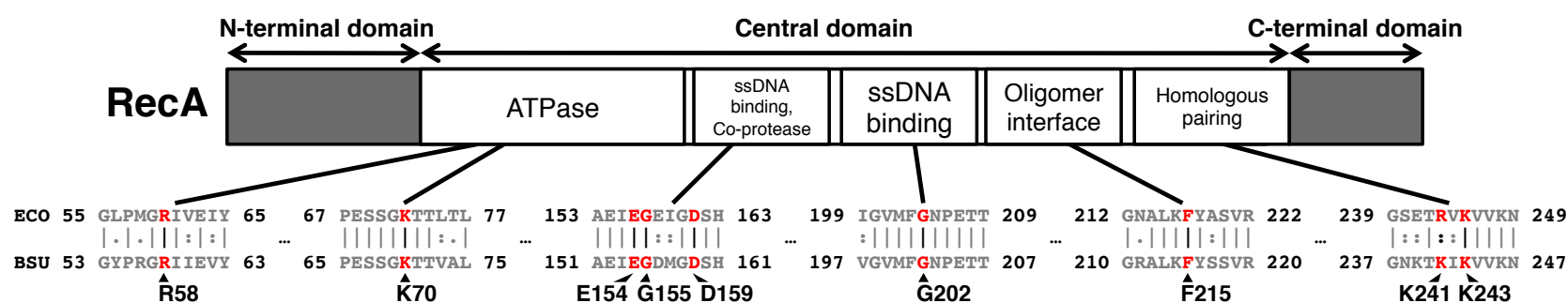
# *recA*変異株を用いたIS256*Bsu1*の 転移頻度の検証

## 4-1. 大腸菌RecAと枯草菌RecAの相同性

ここまでの実験結果から、IS256*Bsu1*トランスポゼースによるmini-ISの転移には*recA*自体が必要であると推察された。では、mini-ISの転移に必要なRecAの機能とは何なのだろうか。一般に、*recA*はこれまで大腸菌にコードされているものが非常に詳細に解析されている。RecAは大別すると3つのドメイン、即ちN-terminal domain, Central domain, C-terminal domainからなる(図5. A)。RecAの主要な活性のうち、既知のものはそのほとんどがCentral domainに集中している。大腸菌のRecAでは、Central domainは53~257アミノ酸の位置に相当する[44]。

まず初めに、EMBOSS stretcherにより枯草菌のRecA(BsRecA)と大腸菌のRecA(EcRecA)についてアミノ酸配列の比較を行った[45]。その結果、両者は非常に高い相同性を有することが分かった(Identity: 57.8% , Similarity: 77.3%, 図5. B)。この結果を元に、大腸菌で知見のあるアミノ酸残基変異株を参照して、枯草菌*recA*活性部位変異株を作製した。

**A**



**B**

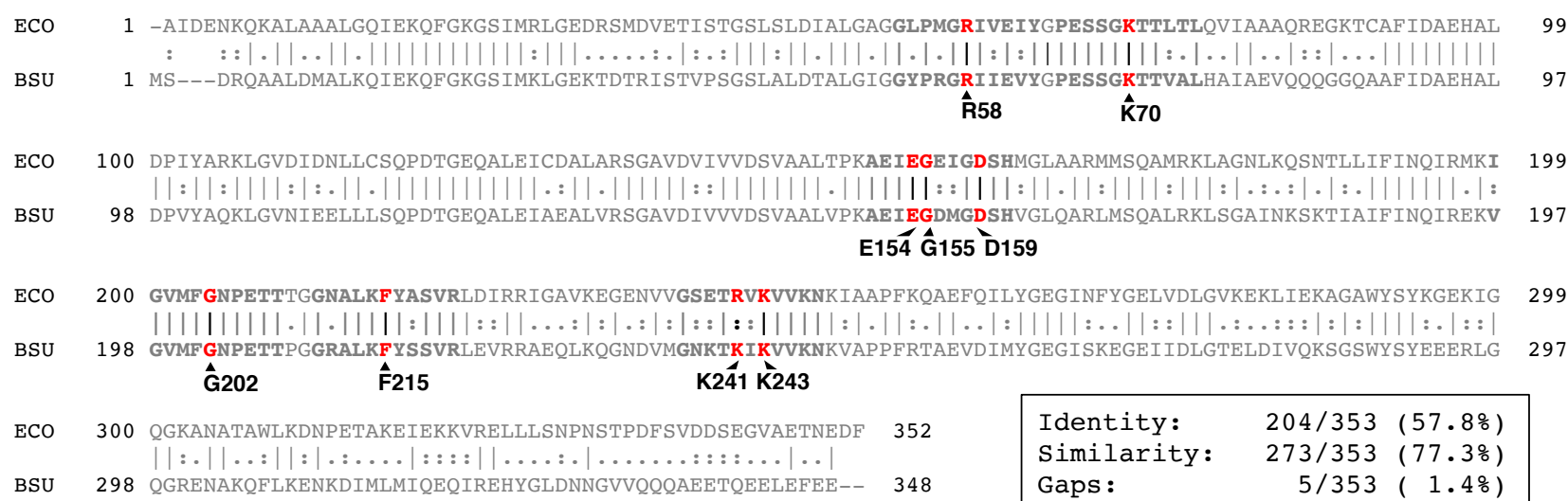


図5. 大腸菌RecAと枯草菌RecAの相同性

**A.** RecAのドメイン構造及び、株作出の際に参照した活性部位を示す（ECO：大腸菌RecA, BSU：枯草菌RecA）。

**B.** EMBOSS stretcherによる大腸菌RecA（ECO）と枯草菌RecA（BSU）のアライメント結果を示す。赤字で示した活性部位は図Aのものと対応している。右下端のボックスは、アライメント結果に基づく相同性を数値化したもの（Identity: アミノ酸が一致した数及び割合, Similarity: アミノ酸の性質が一致した数及び割合, Gaps: アライメントされなかったアミノ酸の数及び割合）。

## <材料, 方法>

### 4-2. 枯草菌*recA*活性部位変異株とmini-IS転移頻度の関係

*recA*変異株は各々, 次の通り作製した. 以下のプライマーの配列及び作製された株の構造は付録を参照のこと. まず, 168株のゲノムを鋳型に, F1-FとF1-R, F2-FとF2-R及びF4-FとF4-Rのプライマを用いて, また, NBS909 (*trpC2 odhB::erm*) のゲノムを鋳型にF3-F, F3-Rのプライマを用いて, PCR反応によりDNA断片を作製した. 続いて, 得られた4種のDNA断片をF1-F, F4-Rのプライマを用いて, リコンビナントPCR法[36]により一つの断片に繋いだものを作製した. 得られたDNA断片を, Jumping cat assay systemを含む枯草菌168株に形質転換し, 薬剤入りの培地で選択, 単離した後, シーケンス解析によりゲノム上の*recA*遺伝子がそれぞれの変異型と置換されていることを確認し, 目的の株を得た. 転移頻度の測定方法は第2章第1節を参照する. Jumping cat assayもこれまでと同様, CI培地における転移頻度を測定, 野生株 (WT) と比較した (図6) .

## <結果, 考察>

作製した枯草菌*recA*活性部位変異株の転移頻度測定結果, 及び考察を, 大腸菌RecAの働きとともに以下に述べる.

### 4-2-1. ATPase活性に関与するアミノ酸残基

ATPase活性に重要な部位における変異体, R58Cと K70Rはそれぞれ大腸菌のR60C及びK72Rを参照して作製した [46,47]. その結果, R58CのTPF野生株に対して約0.198倍まで低下した ( $p < 0.05$ ) . また, K70RからはCm<sup>r</sup>コロニーが生じず, mini-ISの転移が生じていないことが示唆された. したがって, 枯草菌RecAのATPase活性は, mini-ISの転移に極めて重要であることが分かった.

### 4-2-2. ssDNA結合/coprotease/dsDNA結合に関与するアミノ酸残基 (loop L1 region)

loop L1 region の主要な機能は, ssDNA結合及びcoprotease活性であり, この領域にssDNAを結合した状態で*recA*上流のSOS応答配列に結合し, 転写を阻害するLexAと結合,

その加水分解を促す役割を担う[48,49]. 今回の実験で作製した枯草菌*recA*のloop L1 regionの変異体は, E154R / E154V/G155R / G155P / D159Aの5株で, それぞれ大腸菌のE156R / E156V /G157R / G157P / D161Aを参照した[48-50].

転移頻度測定の結果, E154VとG155RにおいてWTよりも3.4倍及び4.3倍それぞれ高い転移頻度を示したが( $p < 0.05$ ), E154RとG155PはWTと比較して統計的に有意な差は見られなかった. Nastriらによると, 大腸菌のE156VはDNA損傷修復 (UV及びMMCによる) は野生株と同等程度で影響が見られない [48]. また, 大腸菌のG157Rは継続的なLexA加水分解活性を示すにも関わらず, DNA損傷修復能は野生株よりも低下する[49]. 枯草菌では, ComK による*recA*の転写は, LexAのSOS応答配列への結合に伴う転写阻害を乗り越えてしまうため[43], もしG155R変異によりLexA加水分解が継続的に行われているならば, *recA*の転写量が増加している可能性は考えられるかもしれない.

近年の大腸菌RecAにおける詳細な解析により, loop L1 regionに位置するD161アミノ酸残基がssDNAまたはdsDNAへの結合性をコントロールしており, この変異体であるD161AはssDNAよりもdsDNAに対して高い結合活性を示すことが明らかとなった[50]. ところが, このD161Aを模して作製した枯草菌RecA D159A変異体は, WTよりも高い転移頻度を示した (2.2倍,  $p < 0.05$ ). つまり, 今回作製したloop L1 regionの変異株は, 野生株よりもmini-ISの転移を促進する傾向にあると言える. 特に, D159A変異株の転移頻度測定結果は, 参照した大腸菌RecA D161Aの性状を考慮すると[50], dsDNAへの結合がmini-ISの転移には重要であると考えられる.

#### 4-2-3. ssDNA結合に関与するアミノ酸残基 (loop L2 region)

大腸菌RecAのloop L2 regionはloop L1 regionとは異なり, ssDNA結合活性だけでなく, 相同なdsDNAとのペアリングも行うことが知られている. このloop L2 regionに変異を持つ大腸菌RecA G204Iは, 相同組換え活性欠損の表現型を示す[51]. そこで, この変異株を参照し, 枯草菌RecAのG202I変異体を作製した. その結果, 転移頻度は, 野生株と比較して7.4倍高い傾向にあることが分かった ( $p = 0.114$ ). つまり, ここでも相同組換え活性はmini-ISの転移には不要であることが示唆されたのである.

#### 4-2-4. RecA間の結合/D-loopの形成に関与するアミノ酸残基

相同組換え反応の際、RecA同士は互いに結合しながらssDNAに螺旋状に結合することが知られており、RecA helical protein filamentと呼ばれる。大腸菌RecAはF217アミノ酸残基を介してRecA helical protein filamentを形成する際のコネクタとして働く [52]。このアミノ酸残基の変異を持つ大腸菌RecA F217Q株は、*in vivo*におけるあらゆるDNA損傷に対する修復能を失っている[53]。この変異株を参照し、枯草菌RecA F215Q株を作製したが、その転移頻度は野性株よりも低下する傾向にあるものの、野性株との間に有意な差は見られなかった(0.6倍,  $p = 0.609$ )。従って、mini-ISの転移に、RecA同士が結合する必要はない可能性が考えられる。

大腸菌RecAのCentral domain のC末端側に位置するR243とK245アミノ酸残基は、RecA-ssDNA複合体 (RecA ssDNA presynaptic nucleoprotein filament) をdsDNAに導入し、D-loopを形成する際に必要なアミノ酸残基である[44,54]。このアミノ酸残基に変異を持つ大腸菌RecA R243Q / K245N二重変異体では、D-loopの形成がされず、然るに相同組換え反応も起きない[44]。このRecA二重変異を有する、枯草菌RecAのK241Q K243N変異体を作製したが、その転移頻度は野性株と変わらなかった(1.3倍,  $p = 0.202$ )。つまり、mini-ISの転移に相同組換え反応のD-loop形成は不要である可能性が高いと考えられる。

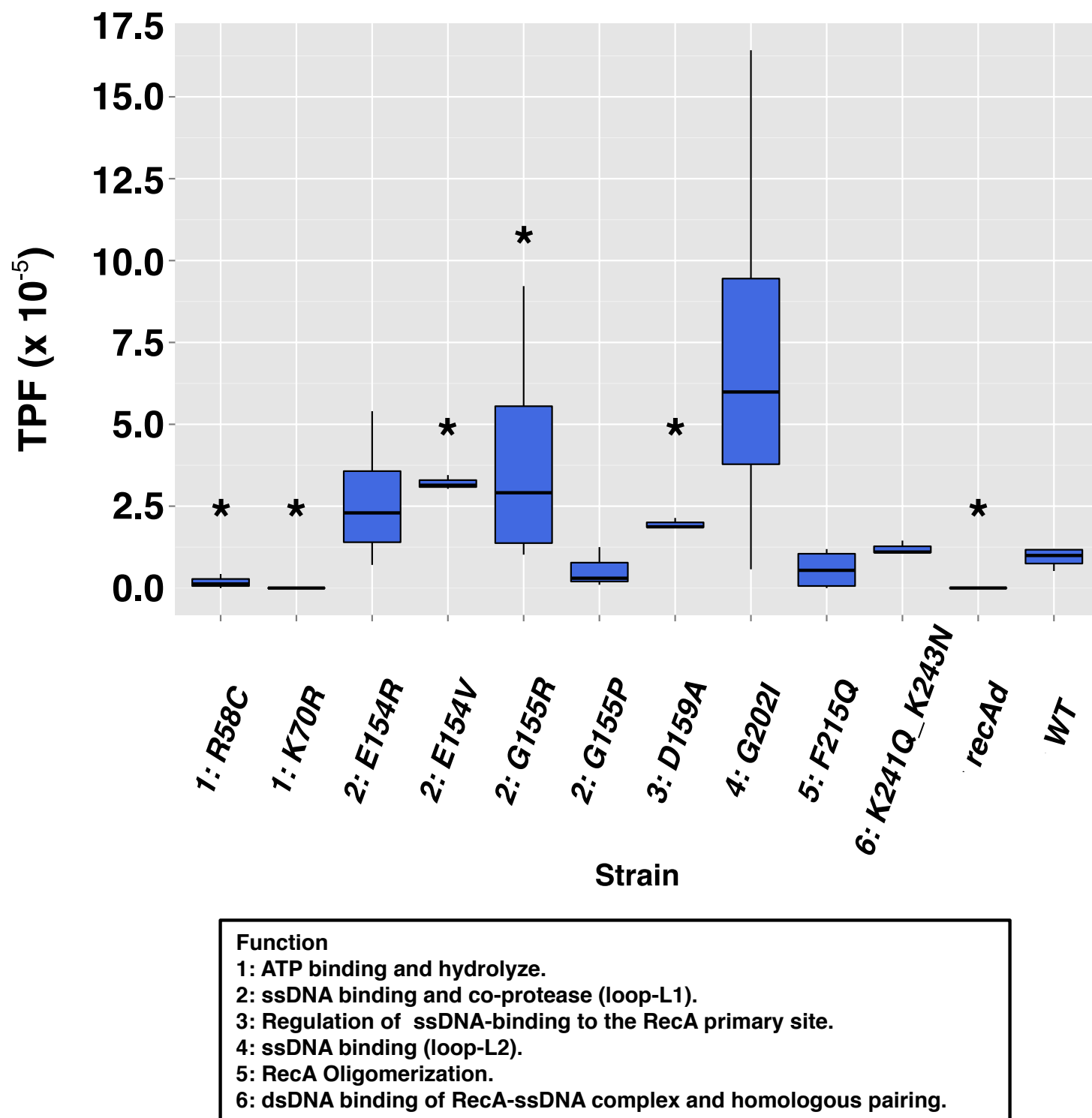


図6. 枯草菌*recA*活性部位変異株とmini-IS転移頻度の関係

mini-IS転移頻度をコンピテンス関連遺伝子欠損株を用いて測定した結果をBox plotで示す。箱の上辺は第一四分位点，中央の太線は中央値，下辺は第三四分位点を示す。箱に対して垂直な直線の上端は最大値，下端は最小値を意味する。TPF（縦軸）は転移頻度，Strain（横軸）はそれぞれ枯草菌RecA上の変異を示す。下端のボックスは各RecA変異株のアミノ酸残基が有する機能を示し，Strainの頭の数字と対応している。アスタリスク（\*）は，並べ替えBrunner-Munzel検定により，CI培地を用いた場合の転移頻度と比較して有意差が認められたものを示す（ $p < 0.05$ ）。尚，参考のため，茂木 俊丞氏の*recAd*の測定値を並べて図示した[27]。

#### 4-3. 小括: *recA*のmini-IS転移に関与する機能, ATPase, dsDNA結合活性

大腸菌と枯草菌の間で, RecAは非常に配列の保存性が高いことが分かったため, 大腸菌RecAのアミノ酸残基変異株の知見を参照しつつ, 枯草菌RecAのmini-IS転移における役割を探った. しかし, これまでの実験と同様, RecAがmini-ISの転移において果たす役割は相同組換えとは異なる機能であることが示唆された. 一方, 次の二つの実験結果, RecAのATPase活性がmini-ISの転移に必須であること, dsDNAへの結合能力が野生型RecAより高い大腸菌RecA D161Aと相同なD159Aのmini-IS転移頻度が野生株と比較して高いことは, 極めて重要な研究上のヒントであると考えられる. つまり, mini-ISの転移には, ATPase活性を持ち, dsDNA結合能を有する因子が必要である可能性がある, ということである.



## 第2章の考察と総括

以上、IS256*Bsu1*の転移に関与する*recA*にコードされた機能について詳細に見てきたが、改めて結果を振り返る。

*recO*と*recU*の欠損株によるJumping cat assayは、相同組換え活性がmini-ISの転移に無関係かどうかを間接的に調べるために行った実験だが、いずれも野生株と比較して顕著な転移頻度の増加が見られたが、これは何故だろうか。やや掘り下げた考察を試みる。枯草菌において、RecOはSSBにコートされたssDNAへとRecAをロードする働きがある[10,55]。RecAの基質の一つであるssDNAへの誘導をRecOが担う為、*recO*が欠損した場合はRecAのdsDNAへの結合確率は上昇し、結果としてmini-ISの転移頻度も上昇すると考えられる。RecUは、コンピテンスにおいて、RecAとともに細胞の極に局在し、取り込まれるssDNAを待ち構えており、*recU*が欠損すると染色体の形質転換効率も半分近くまで低下する[11]。このことは、RecUが細胞内に存在しなければ、DNAの取り込みや、その後の相同組換えに関わる反応の誘導される割合も恐らく低下することを示唆しており、結果としてRecAのdsDNAへの結合する割合を上昇させると考えられる。これが、筆者の考える*recU*欠損によりmini-ISの転移頻度が上昇する主な要因である。また、これは推測の域を出ないが、RecUはリゾルベースであることから、IS256*Bsu1*によるmini-ISの転移反応を直接阻害しており、欠損によりその阻害効果が解消される可能性も考えられる。

*recA* D159Aにおける結果は、他の*recA*変異株の内、転移頻度が有意に上昇したものの(E154V, G155R, G202I)についても説明することが可能である。大腸菌RecAのD161に関する知見は、言い換えれば、RecAのD161を介して、ssDNAに結合して相同鎖探索を行うモードとdsDNAに結合するモードが切り替えられているということである[50]。従って、RecAは酵素的に、二種類の基質をターゲットにすることになる。E154V, G155R, G202Iの三つのアミノ酸残基はいずれもRecAのssDNAに対する結合性に関与する残基である[48,49]。これらのアミノ酸残基がssDNAを捉える確率が低下するならば、自動的にdsDNAへの結合確率が上昇すると考えられる。例えば、大腸菌RecAのD161A変異は、野生株よりもdsDNAへ結合し

やすく、野生型のRecAの相同DNA結合の反応速度はssDNAと結合した状態でdsDNAと混合した時に最も大きく、他の組合せでは小さくなるが、D161Aは反応速度も基質の種類や状態によらず野生型のRecAにおける最大速度と同程度である[50]。このことは、通常の野生型RecAはdsDNAでなくssDNAに結合しやすく、ssDNAへの結合能が低下した場合には他の基質であるdsDNAに結合する確率が上昇するということを意味する。以上を踏まえると、枯草菌RecAのE154V, G155R, G202I変異株において転移頻度が上昇したことは、RecAのdsDNA結合性がmini-ISの転移に必要なことを間接的に示唆していると考えられる。

また、注意したいのは、枯草菌RecAのK241Q K243Nに相同な大腸菌RecA R243Q K245N変異に基づくこれらのアミノ酸残基の働きは、ssDNA-RecA 複合体の状態でのdsDNAの相補鎖探索であり、RecAそのもののdsDNAへの結合ではないことである[44,54]。枯草菌RecA K241Q K243Nにおける転移頻度が野生株と変わらないのは、ssDNAとRecAが複合体を形成した状態でdsDNAに滑り込み、相補鎖を探索するというプロセスが、mini-ISの転移には全く必要がなく、しかもこれらの変異がdsDNAへのRecAの結合を阻害しない為であると考えられる。

# 第3章 本論-2

バクテリオファージMuにコードされる *muB* 遺伝子と

相同組換え酵素遺伝子の分子系統学的関係性,

およびトランスポゾンの転移における *muB* 遺伝子の

相同組換え酵素遺伝子に対する相補性について.

# バクテリオファージMuにコードされた

## *muB*と*recA*を含む

## 相同組換え因子の系統解析

本節では、研究の過程で*muB*に着目した理由について述べる。また、*recA*, *radA*, *rad51*, *uvsX*など、*recA*と共通祖先配列を有すると考えられている遺伝子と*muB*の系統関係を調査したので、その方法及び結果を記載する。また、系統関係と*recA*及び*radA*遺伝子の塩基配列のGC含量について観察された相関性についても述べる。

### 1-1. IS256*Bsu1*と*muA*は共にDDEタイプの転移酵素をコードする

ここまでの遺伝学的実験の結果から、IS256*Bsu1*のmini-ISの転移には、*recA*遺伝子は必要だが、相同組換えに関与する機能が保存されている必要はないようである。では、なぜ*recA*はmini-ISの転移に必要なのだろうか。この問題を考える上で、Enterobacteria phage Muのゲノム内における転移酵素MuAに着目した。挿入配列 (IS) と呼ばれるトランスポゾンにコードされる転移酵素は、活性部位であるアスパラギン酸とグルタミン酸 (DDEモチーフ) を除くと配列の保存性が低い。そのため、分子系統解析が極めて困難であるが、DDEタイプの転移酵素について、近年の研究によりその進化的な関係性が徐々に明らかになりつつある。本実験に用いたIS256*Bsu1*の属すIS256 familyは、prokaryotic Mutator-Like transposaseに属するDDEタイプの転移酵素(p-MULT1)である[22]。また、真核生物のコードするDDE/D型のトランスポゾンは共通の祖先が存在すること[23]、トウモロコシ (*Zea mays*) にコードされている*MuRA* (Mutator transposaseの一つ) とIS256の配列には保存性があることから[24]、IS256 familyは真核生物にも見られる転移酵素と共通祖先を持つことが分かる。また、真核生物とp-MULTのアミノ酸配列の二次構造予測を比較した結果は、いずれもそのモチー

フが特異的に保存されていることを示している[22]. 更に, MuファージにコードされているMuAもDDEタイプの転移酵素であり, これまで詳細に解析されてきた[25]. 従って, DDEタイプの転移酵素に共通祖先が存在することを示唆しており, 結果としてIS256Bsu1とmuAの間に遺伝的な関係性があると考えられる.

### 1-2. RecAとMuBの共通点

Enterobacteria phage Muにコードされ, RecAと同じP-loop dNTPase superfamily に属すDNA結合タンパク質MuBは, MuAの転移反応を補佐することが知られているが[56], MuAとIS256Bsu1の場合と同様, MuBとRecAには数多くの性質上の共通点を見出すことができる. MuBはATPase活性を有し, ATPと結合した状態でDNAに結合し, DNA-MuA複合体をターゲットサイトに呼び込んだ後, ATPの加水分解を伴ってDNAから解離する[19,57-59]. RecAにはssDNAまたはdsDNAと結合することの出来るサイトが存在し, dsDNAから解離する際にはATPの加水分解を伴うとされる[50,60,61]. つまり, 両者の共通点をまとめると, P-loop dNTPase superfamilyに属すること, ATPを結合した状態でdsDNAと結合すること, かつDNAから解離する際にATPを加水分解することが挙げられる.

### 1-3. MuBとRecAアミノ酸配列の類似性

まず, MuBと枯草菌のRecAの比較を行った. 同時に, 真核生物のRecAホモログであるRad51をRecAと比較し, MuB-RecA間の相同性との比較も行った. 配列の比較はアミノ酸配列を用い, 比較するプログラムにはEMBOSS Stretcher[45]を使用した. まず, MuBのアミノ酸配列(NP\_050608.1)と枯草菌におけるRecAのアミノ酸配列(WP\_003245789.1)の配列の相同性は, Identity 17.7%, Similarity 34.5%であった. 一方, 枯草菌RecAと酵母のRad51(NP\_011021)はIdentity: 16.1%, Similarity: 28.8%, 枯草菌RecAとHomo sapiensのRad51(NP\_002866.2)はIdentity: 13.1%, Similarity: 29.5%であった. つまり, 配列の類似性の点から考えるならば, 互いに相同な因子であるRecAとRad51よりも, RecAとMuBの配列は似ているということになる. 従って, recAとrad51が進化的に共通な祖先配列を有するならば, muBもこの祖先配列を共有している可能性がある.

Identity / Similarity (%)	BsRecA	MuB	ScRad51	HsRad51
<b>BsRecA</b>				
<b>MuB</b>	17.7 / 34.5			
<b>ScRad51</b>	16.1 / 28.8	16.1 / 32.8		
<b>HsRad51</b>	13.1 / 29.5	16.4 / 33.4	57.0 / 70.0	

**表2. MuBとRecAアミノ酸配列の類似性**

枯草菌のRecA (BsRecA, WP\_003245789.1) , 酵母のRad51 (ScRad51, NP\_011021) , *Homo sapiens*のRad51 (HsRad51, NP\_002866.2) を MuBのアミノ酸配列 (NP\_050608.1) とEMBOSS Stretcherで比較した結果を示す. Identityはアミノ酸が一致した割合 (%), Similarityはアミノ酸の性質が一致した割合 (%) を示す.

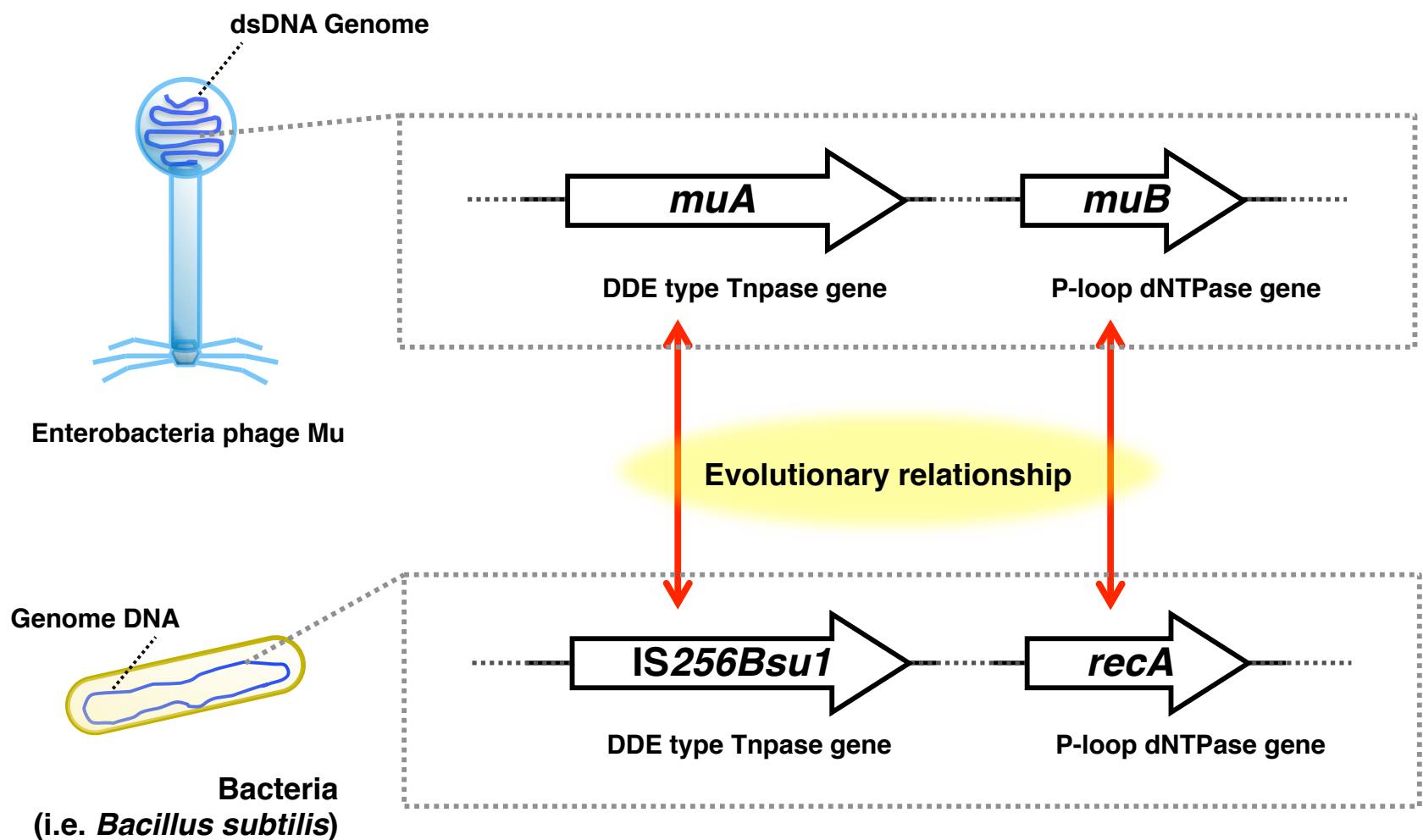


図7. *recA*と*muB*, *IS256Bsu1*と*muA*に見られる類似性

*muA*と*IS256Bsu1*はDDE typeのトランスポゼース (Tnpase) をコードする遺伝子, *muB*と*recA*はP-loop dNTPaseをコードする遺伝子である. *muA*と*muB*はMuファージ (Enterobacteria phage Mu) の二本鎖DNAゲノム (dsDNA Genome) 上に存在する.

#### 1-4. *recA/rad51/radA/uvrX*及び*muB*の分子系統解析

以上から、MuBの他の相同組換え因子 (RecA like superfamily) との系統関係の調査を行った。系統解析には、バクテリアにコードされる*recA*だけでなく、他の主要な相同組換え・DNA修復因子である*rad51*, *radA*, *uvrX*も加え、*muB*遺伝子の生物界における系統的な位置付けを明らかにすることを目指した。*rad51*は真核生物、*radA*は古細菌、*uvrX*はT4ファージの仲間がそれぞれ有する相同組換え因子であり、いずれもペプチド配列の保存性が指摘されている [62,63]。なお、これらの遺伝子の塩基配列を全ゲノム解析が進行中、あるいは完了した生物種より収集した。植物はミトコンドリア及び葉緑体にコードされる*recA*を必ず保存していたので、*rad51*とともにデータに加えた。また、細胞性粘菌 (Mycetozoa) も*rad51*だけでなく*recA*をコードしていることが知られており[64]、これもデータに加えた。全てのシーケンスデータはORFに間違いが無いか一つずつ確認した。極力、既知の生物種が網羅的に含まれるように配列を選び出したが、最終的に解析に使用したOTU (operational taxonomic unit) は406であった。なお、配列はNCBI genome databaseより取得した。解析に用いた配列番号は、生物種名とともに付録2に記載した。

#### <方法>

分子系統解析の手順は次の通りである。まず初めに、全ての塩基配列をペプチド配列へ変換した。この配列群を、遺伝子の種類ごと (*recA*, *rad51*, *radA*, *uvrX*, *muB*の5通り) に分類し、分類群ごとにMAFFT (version 7.273, “maxiterate1000”, “genafpair”オプションを指定) を用いて整列し、特異的に挿入されたと見られる配列を除去した[65]。続いて、余分な挿入を除去した配列群を一つにまとめ、改めてMAFFT (“maxiterate1000”, “localpair”オプションを指定) を用いて整列した[66]。その後、整列された配列の内、アライメントの精度が低い領域を、TCS (Transitive consistency score, T-Coffee version.11.00.8cbe486で算出, “filter3”オプションを使用) に基づき除去し、同時にペプチド配列となっているデータをDNA配列に戻した[66]。相同組換え因子のアミノ酸配列は、各々非常に保存性が高いが、これをDNA配列に戻すことにより、配列上のアミノ酸に対応する様々なコドンと比較でき、結果として情報量が増えるため、精度の高い系統樹が得られる。分子系統樹の推定にはGeneral Time Reversible modelに基づく最尤法 (ML法, Maximum Likelihood method) を使用した



[67]. また, bootstrap法により系統樹の各分岐点の信頼性を計算し (試行回数は200回とした) [68], 得られた系統樹の枝の根元に枝の再現性をパーセントイルで示した. 最終的に1,044bpの座位が系統推定に使用された. なお, 分子系統樹の作製にはMEGA7 (version 7.0.14) を使用した[69]. また, 系統樹の描画にはFigTree (version 1.4.2) を使用した (<http://tree.bio.ed.ac.uk/software/figtree/>) .

## <結果, 考察>

系統解析により得られた系統樹を図8に示す. OTUは全てClassごとにまとめた. 系統樹は, 解析に用いた5タイプの遺伝子 (*recA*, *rad51*, *radA*, *UvsX*, *muB*) からなる5つの巨大クレードが形成された. まず, *muB*遺伝子の他の相同組換え遺伝子の中での系統的な位置付けを確認した. *muB*と*UvsX*はいずれもバクテリオファージに由来する遺伝子であるが, *muB*のクレードは*UvsX*のクレードと隣り合った位置に存在し, *UvsX*や他の相同組換え因子をコードする遺伝子のクレードと混ざらず, 独立して存在する(>99%). 従って, *muB*は*recA*や他の相同組換え因子を元に, 後から派生した遺伝子ではないことが示唆された. 他の遺伝子群の関係についても, *muB*と同様に確認を行った. *rad51*のOTUは全て真核生物によって構成されているが, このクレードは*radA*のクレードと隣り合った位置で互いに分離された(>98%). *recA*のクレードは植物の葉緑体, ミトコンドリアの*recA*を含むクレードを除くと, ほぼ全てバクテリアにより構成されたクレードであるが, そのOTUの構成は不調和的であり, 同じ綱 (class) のクレードが異なる位置に見られる. この傾向は*radA*のクレードにも見られた. 古細菌の一種である*Halosimplex carlsbadense* 2-9-1 (NZ\_AOIU01000018.1: 12005–13468 bp)に由来する「recombinase RecA」とアノテーションされた配列に基づく枝は, 他の全てのクレードから独立して存在した. 元となる配列を確認したところ, 二つの重複したATPaseドメインが存在したため, この配列が実際には「*kaiC*」である可能性が高い[63]. 実際, *H. carlsbadense* 2-9-1の*radA*はhalobacteriaのクレードに分類されている. なお, 枝を束ねていない系統樹を付録に示す.

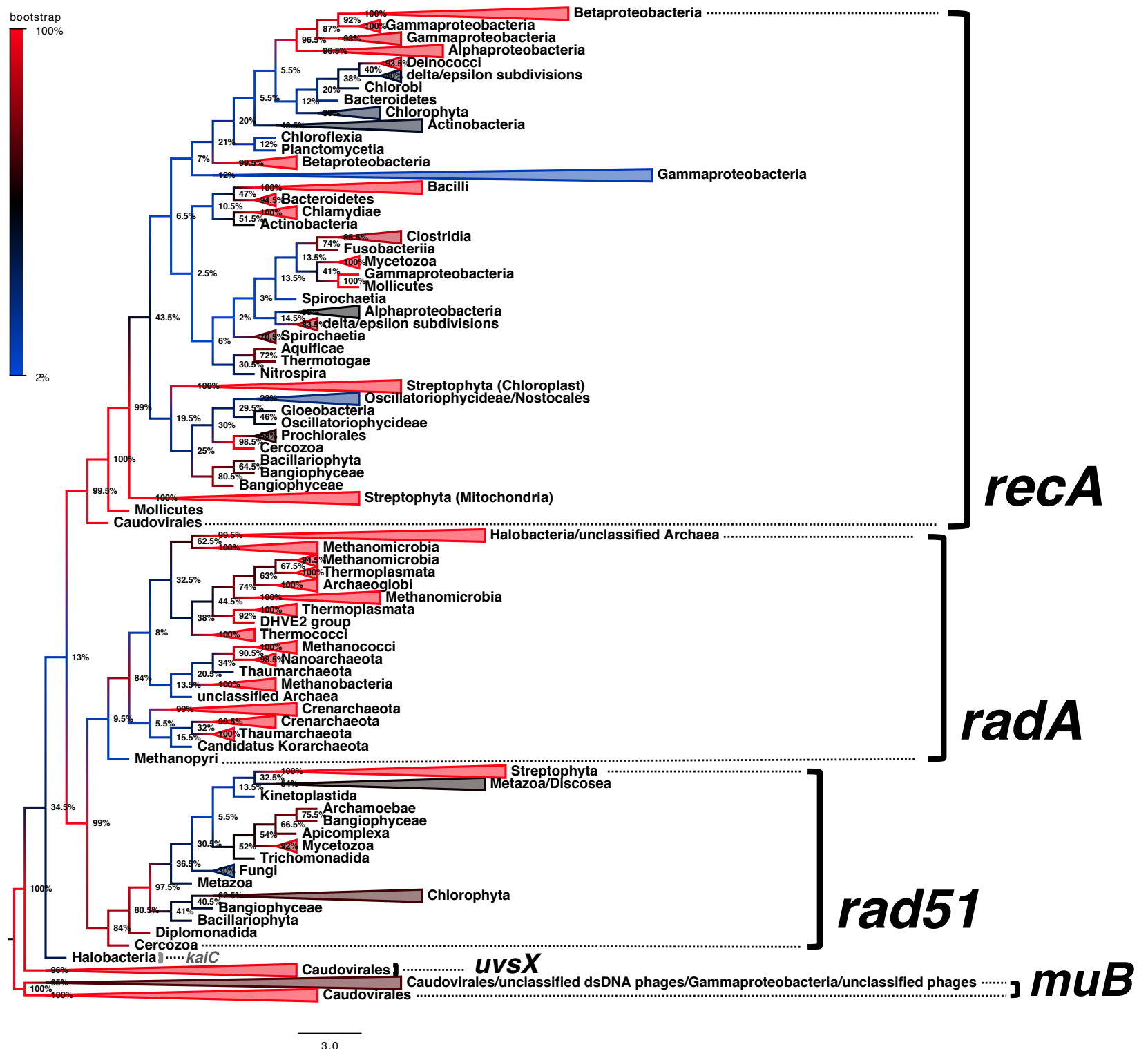


図8. *recA*/*rad51*/*radA*/*uvsX*及び*muB*の分子系統解析

General Time Reversible modelに基づく最尤法の分子系統樹（bootstrap-consensus phylogenetic tree）を示す。カラスケールは、分岐のBootstrap確率に対応している。枝はそれぞれ網ごとにまとめ、その名称を記載した。枝を纏めていない系統樹は付録3に記載した。

## 1-5. *recA*及び*radA*の塩基配列が示す系統関係とGC含量の相関性について

ファージやplasmidの変異傾向はグアニン (G) やシトシン (C) に偏向していることから、宿主のゲノムのGC含量は、ファージや水平伝播するプラスミドの感染によって徐々に増加すると考えられている [70]. そこで、得られた系統樹を用い、OTU間の系統関係と各塩基配列のGC含量の相関性についても調査した。

### <方法, 結果, 考察>

図9. Dに矢印で示した*recA*及び*radA*のクレードについて、系統樹の枝の位置に基付き、系統解析に用いた配列のGC含量をプロットすると、それぞれ線型にGC含量が増加していることが分かった ( $R^2 = 0.81$ 及び $0.79$ , 図9. A/B). つまり、この*recA*と*radA*のクレードでは、枝の分岐回数が増えるに従って、配列のGC含量が増加するということである。では、解析に使用した遺伝子のGC含量と、その遺伝子を持つ生物の全ゲノムのGC含量には相関性が見られるのだろうか。実際に、今回解析に使用した生物のゲノムGC含量に対する各遺伝子のGC含量をプロットした。各生物のゲノムGC含量はNCBI Genome Databaseを参照した。その結果、これらの値は互いに高い相関性を示すが、全体的にゲノムのGC含量よりも相同組換え遺伝子のGC含量の方が高い傾向にあり、ゲノムGC含量が高い生物の持つ相同組換え遺伝子ほど、そのGC含量がゲノムの値に近いことが分かった ( $R^2 = 0.89$ , 図9. C). このような分子系統解析と解析に用いた塩基配列のGC含量の関係性が偶然に生じることは考えられない。何故ならば、あるクレード内にある配列の子ノードに配置された配列は、一つ上の階層に配置された配列と類似していることは意味するが、塩基配列の塩基の種類に影響されないためである。つまり、一般的に、ある子ノードに接続した子ノードのGC含量が高いという必然性はない。実際、古細菌とバクテリアに見られた相関性は、他のクレードでは見られない (詳細は付録4を参照のこと)。以上を踏まえると、バクテリア及び古細菌の相同組換え遺伝子の系統関係とGC含量の間に見られる相関性は、*recA*や*radA*の配列上のアデニン (A) やチミン (T) がグアニン (G) やシトシン (C) へ、まるで宿主のゲノムのGC含量に近づくように変異していたことを示唆する。以上から、*recA*や*rad51*, *radA*は*muB*や*uvsX*と共通の祖先配列に由来し、水平伝播を介して獲得された遺伝子であることが示唆された。

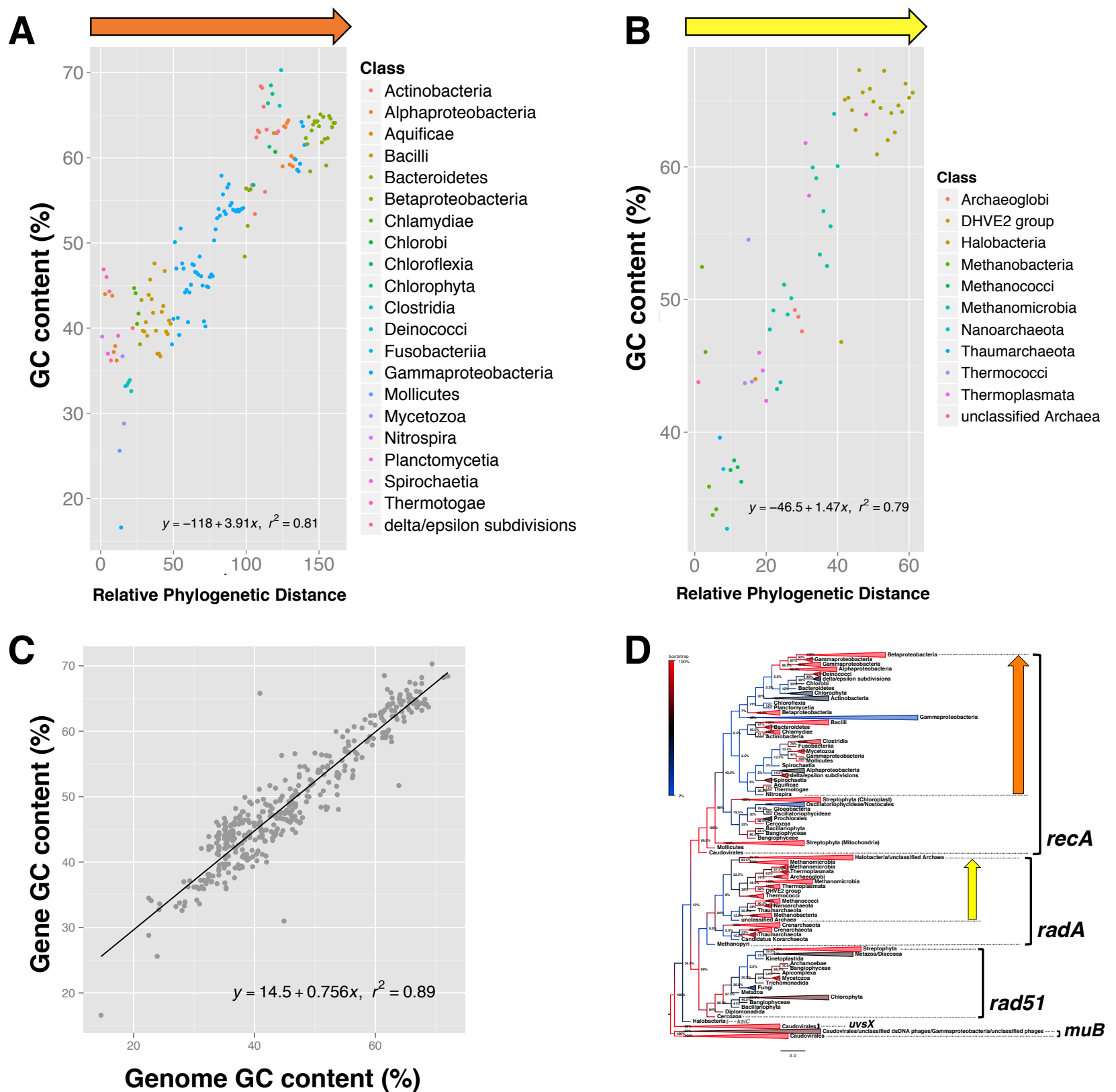


図9. *recA*及び*radA*の塩基配列が示す系統関係とGC含量の相関性について

**A, B.** バクテリア及び古細菌における、相対的な系統関係に対する各遺伝子のGC含量をプロットした図を示す。系統樹における矢印で示した該当箇所（図D）が、図A, Bにおけるオレンジ及び黄色の矢印として対応している。各図右下の数式は線型近似した時の関数、 $r^2$ は相関係数を示す。

**C.** 解析した全ての種における、相同組換え遺伝子に対する、そのゲノムのGC含量をプロットした図を示す。右下の数式は線型近似した時の関数、 $r^2$ は相関係数を示す。

# IS256Bsu1の転移における *muB*と*recA*の相補性の検証

mini-ISの転移に関して、*muB*が*recA*を相補するかについて述べる。同時に、紫外線（UV）照射に対しても*muB*が*recA*を相補するか検証した結果も合わせて記載する。

## 2-1. *muB*遺伝子のmini-ISの転移における*recA*への相補性

以上の研究を踏まえ、*muB*が、IS256Bsu1によるmini-ISの転移における、*recA*の必須性を相補できるかを検証した。*muA*と*muB*は互いにMu フェージ上の遺伝子であり、*muA*の転移反応には*muB*が必要である。*muA*とIS256Bsu1はともにDDE型の転移酵素である。系統解析の結果より、*muB*と*recA*には共通の祖先配列が存在したと考えられる。従って、IS256Bsu1の転移には*recA*が必要であるという関係性が、*muA*と*muB*の関係と同じならば、IS256Bsu1の転移に対する*recA*の必須性は、*muB*により補完されると考えられる。

### <材料, 方法>

そこでまず、*muB*のORFを、*recA*のORFと完全に置換したJumping cat assayシステムを有する枯草菌を作製した。*muB*遺伝子は*Escherichia coli* ATCC 9637に溶原化しているMu フェージより取得した（CP002967.1, 325221- 326159bp, ORFの向きは逆向き）。つまり、この株は*recA*の代わりに*muB*をコードしていることになる（*recA*のORFは一切含まない）。具体的な*recA::muB*置換株の作製方法は次の通りである。以下のプライマーの配列及び作製された株の構造は付録を参照のこと。まず、次の三通りの組み合わせでPCR反応を行い、3種類のDNA断片を得た（枯草菌168株のゲノムとF1-F, F1-Rのプライマ、*E.coli* ATCC 9637株のゲノムとF2-F, F2-R, NBS290 (*trpC2 recA::tet*) のゲノムとF3-F, F3-Rのプライマ）。続いて、得られた3種類のDNA断片を、F1-F, F3-Rのプライマを用いて、リコンビナントPCR法

[36]により一つの断片に繋いだものを作製した。得られたDNA断片を、Jumping cat assay systemを含む枯草菌168株に形質転換し、薬剤入りの培地で選択、単離後、シーケンス解析により*muB*遺伝子が*recA*遺伝子と置換されていることを確認し、目的の株を得た。

## <結果>

この株を用いて、これまで通りCI培地を用いてJumping cat assayを行った（転移頻度の測定方法は第2章第1節を参照する）。その結果、*recA::muB*置換株の転移頻度は、*recA*を持たないにも拘らず、*recA*を有する場合（つまりWT）と同等であり、互いに統計的なレベルで差は無いことが分かった（0.93倍、 $p = 0.40$ 、図10）。つまり、IS256*Bsu1*によるmini-ISの転移に対する*recA*の必須性は、*muB*により相補されることが分かった。

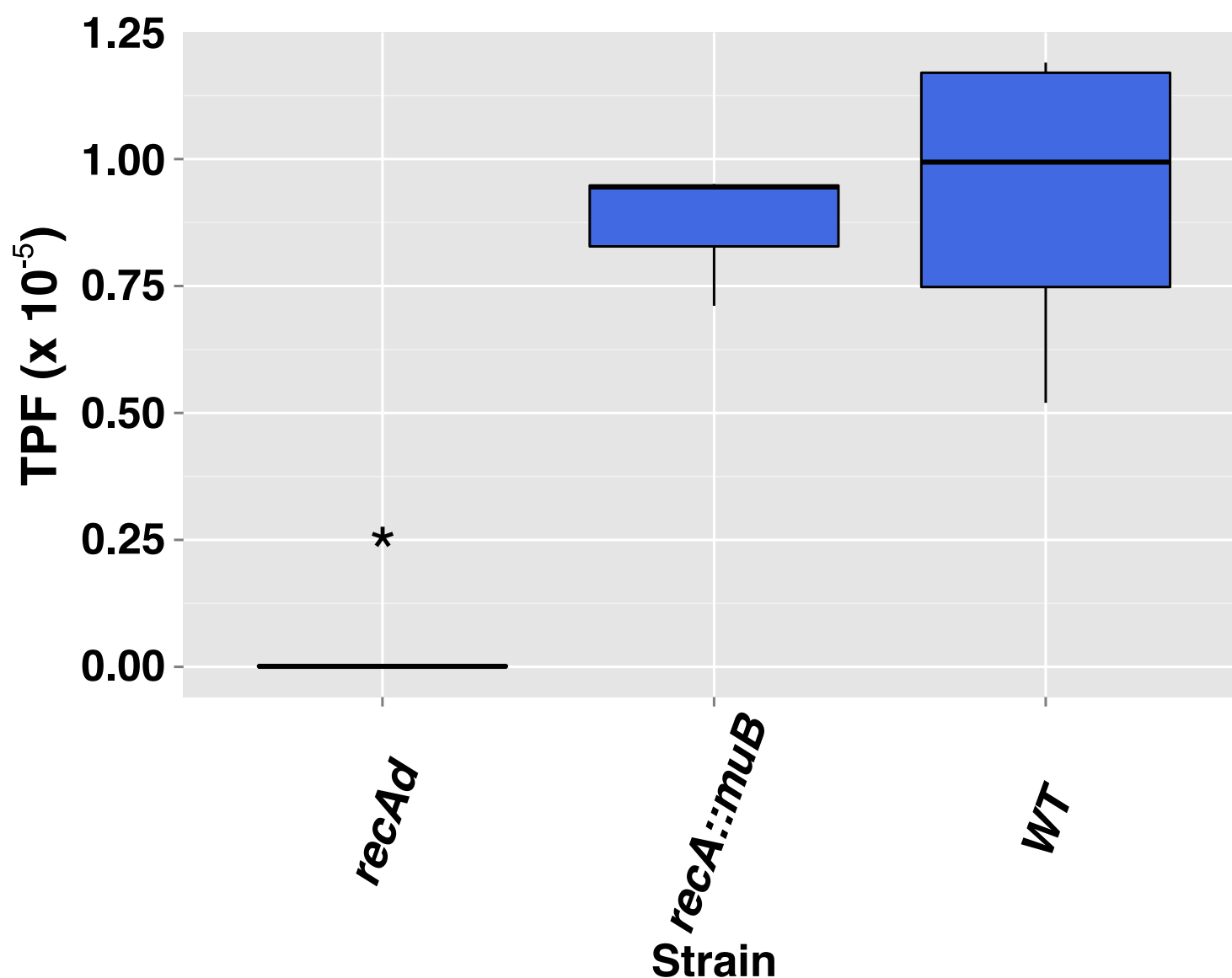


図10. *muB*遺伝子のmini-ISの転移における*recA*への相補性

mini-IS転移頻度を*recA::muB*置換株を用いて測定した結果をBox plotで示す。箱の上辺は第一四分位点，中央の太線は中央値，下辺は第三四分位点を示す。箱に対して垂直な直線の上端は最大値，下端は最小値を意味する。TPF（縦軸）は転移頻度，Strain（横軸）は各株の遺伝型を示す（*recAd*：*recA*欠損，*recA::muB*：*recA*遺伝子を*muB*遺伝子で置換した株，WT：野生株）。アスタリスク（\*）は，並べ替えBrunner-Munzel検定により，CI培地を用いた場合の転移頻度と比較して有意差が認められたものを示す（ $p < 0.05$ ）。尚，参考のため，茂木 俊丞氏の*recAd*の測定値を並べて図示した[27]。

## 2-2. *muB*遺伝子のUV耐性における*recA*への相補性

続いて、これらの*muB*の*recA*に対する相補性が、mini-ISの転移に限定されるのか、それともDNA修復能も相補可能かを調べるため、*recA::muB*置換株のUVに対する耐性を検証した。

### <材料, 方法>

実験は、半定量的な大腸菌のUV耐性検証法として過去に報告のある方法を参照した[70]. 実験には通常の枯草菌168株, Jumping cat assayシステムを有する枯草菌168株3種 (*WT*,  $\Delta$ *recA*, *recA::muB*) を用いた. 凍結保存したこれらの株は, LB寒天培地に塗り広げて一晚培養し, 翌日LB液体培地で吸光度 (600nm) が1.0となるまで培養した. この培養液を, 白金耳でLB寒天培地に直線状に塗り広げた. プレートに光の反射率が小さい硬質ウレタンで部分的に覆いながら, 照射線量 ( $\text{J}/\text{m}^2$ ) が昇順となるようUV-Cを照射した. その後, 37°Cで16時間, 光を遮蔽して培養し, *recA::muB*置換株のUV耐性を他の株と比較した.

### <結果, 考察>

実験の結果, *muB*と*recA*を置換した株のUV耐性は, *recA*を欠損している場合と同様, 非常に低いことが分かった (図11). また, 野生株について, Jumping cat assayシステムを含むものとそうでないものを比較し, Jumping cat assayシステムの有無はUV耐性とは無関係であることも分かった (図11). 以上を踏まえると, *recA*は, IS256*Bsu1*によるmini-ISの転移に寄与するが, その役割は少なくとも*in vivo*では相同組換え反応とは異なる機能であり, かつ*muB*に保存されているものと同様である. つまり, IS256*Bsu1*を介した*muB*と*recA*の系統的なつながりの存在が示唆された. 一方, *recA*や他の相同組換え因子に見られるDNA修復能は, *muB*と*recA*の共通祖先配列が変異する中で獲得されたものであると考えられる.



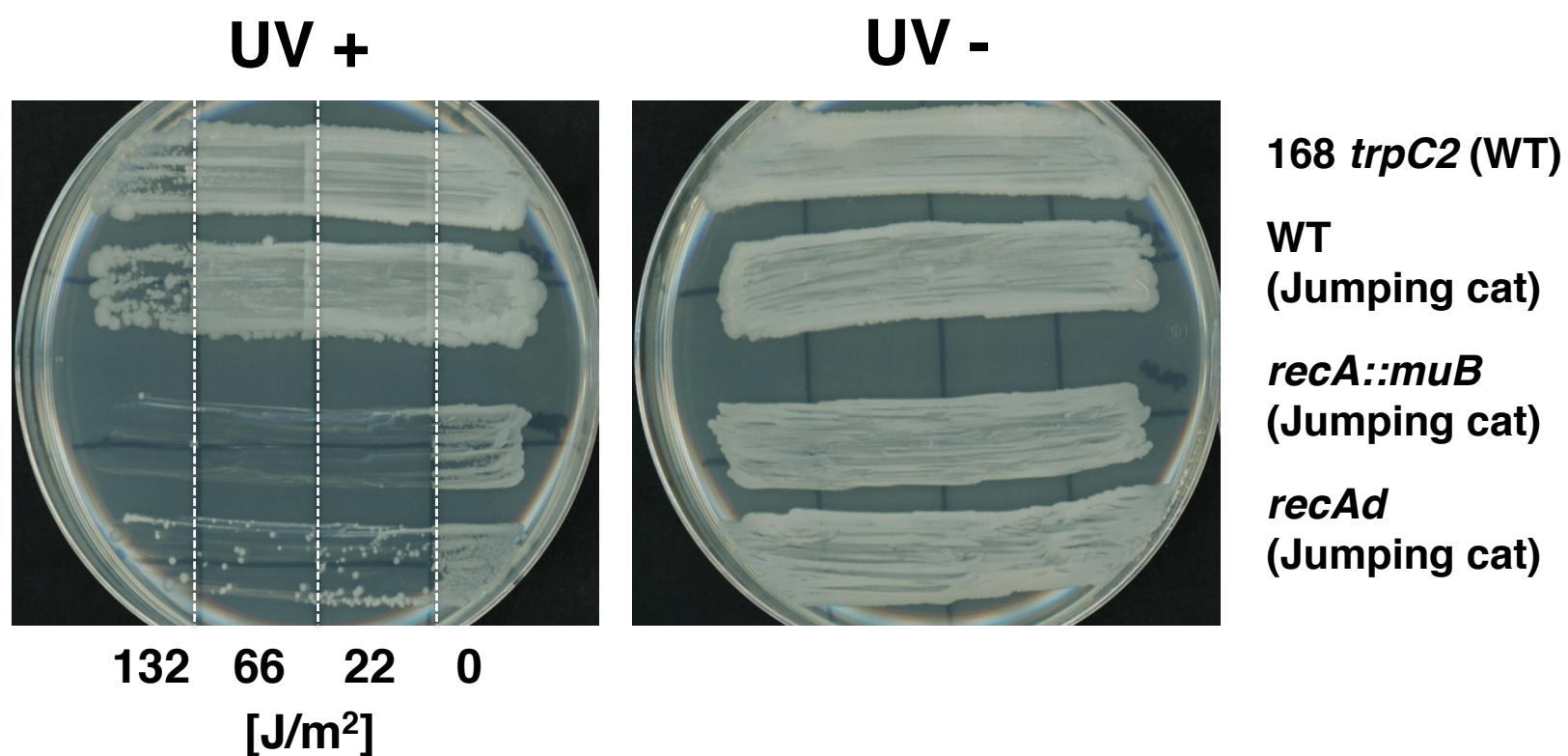


図11. UV耐性における*muB*遺伝子の*recA*に対する相補性試験

*recA::muB*置換株の，*recA*におけるUV耐性の相補性を検証した結果を示す．左の写真はUV照射群（UV+），右はUV非照射群（UV-），右端は塗抹した菌株の遺伝型に対応する．UV照射群（UV+）における縦線は，照射した紫外線の強度を区別するためのものであり，各々の紫外線の照射強度を写真下に記載した(J/m<sup>2</sup>)．

## 第3章の考察と総括

以上, IS256Bsu1の転移における*recA*と*muB*の関係について見てきた. 系統解析の結果は, 相同組換え酵素がバクテリオファージMuやT4-like virusなどと共通祖先配列を有し, しかも進化的には相同組換え酵素遺伝子がバクテリオファージなどのウイルスに由来することを示唆している. また, *muB*と*recA*がIS256Bsu1によるmini-ISの転移において相補的な関係にあることは, *muB*と*recA*にコードされる因子が, 非常にマクロなスケールで互いに共通した機能を保持していることを意味する. 実際のところ, これはどのような機能なのだろうか. 主要なMuBの働きはMuA-DNA複合体を標的領域に誘導することであり, ATPを加水分解してdsDNAから解離する[19,57-59]. RecAも同様にdsDNAに結合し (ssDNA結合性は共通する活性ではない), そしてATP加水分解を伴って, dsDNAから解離する[50,60,61]. RecAとMuBのアミノ酸配列が相同性が比較的高いことは既に述べた通りである. MuB配列を基準として配列に関する違いを確認すると, 主にはRecA配列にMuBにはない配列が挿入されていることが分かる. MuBに存在する3つのドメインのうち, N-terminal appendageドメインと $\alpha/\beta$ ドメインの間 (65-77 aa) の区間は, RecAの配列における相同な領域の間に12 aaの挿入が見られた (図12). また, MuBのLoop 1領域と相同なRecAの領域の直後にも, 16 aaの挿入が見られるが, その挿入前後のアミノ酸は互いに高い相同性が見られる (MuB: 156-SRALRR-161, BsRecA: 170-SQALRK-175, 図12). MuBのLoop 1における塩基性残基に対する変異 (MuB R150A/R151G/K152A, R151A/K152A) はMuAによるATPase活性の活性化に応答しない[72]. 今回実際に変異株を作製した枯草菌RecAのアミノ酸残基とMuBのドメインとの対応関係を調査すると, 枯草菌RecAのR58はN-terminal appendageドメイン, E154, G155, D159, G202はMuBにおけるAAA+ module内の $\alpha/\beta$ ドメイン上に存在した. 特に, 枯草菌RecA G202は, MuBの $\alpha/\beta$ ドメイン内におけるWalker B motifと相同な領域に存在した. また, K70, F215, K241, K243はMuBに相同な残基は見られないが, その前後の領域の保存性は高い (図12). MuBのN-terminal appendageドメインは, ATP加水分解, MuAとの相互作用, そしてDNA結合のいずれにも関係が無い[72]. 一方で, 大腸菌K72R変異株が示すように[47], RecAのCentral domain N末端側のATPase活性に関与する領域はATP加水分解

活性において重要である。また、筆者の行った実験においても、この残基の変異 (K70R) ではmini-ISの転移は検出出来なかった (図6)。このことから、ATP加水分解活性に関わる残基は、MuBとRecAでは異なると考えられる。ただし、IS256Bsu1の転移と、RecAのATPase活性の必要性は、MuAの転移反応におけるMuBのATPase活性の必要性と同じなのでは無いだろうか。Mizunoらのモデルを参照すると、MuAはMuB-ATPの加水分解を促進することにより、dsDNA上にフィラメントを形成していたMuBが一部解離する。MuBとMuAの相互作用はMuの末端にニックを入れると同時に、MuAによる露出したdsDNA上への転移反応が開始される[72]。おそらく、RecAがIS256Bsu1の転移に介在する場合、このモデルと同様、トランスポゼース-mini-IS複合体をRecAがターゲット領域に呼び込み、ATP加水分解を伴ってDNAから解離した後、トランスポゼースがmini-ISの両端にニックを入れ、ターゲット領域への転移を促すと考えられる。大腸菌RecAのD161はloop L1領域に存在し、ssDNAに結合を伴うdsDNA上の相同鎖の探索、またはdsDNAへの結合のいずれかのモードを決定する際に重要なアミノ酸残基であると考えられている[50]。このD161と相同なMuBのアミノ酸残基は $\alpha/\beta$ ドメインのLoop 1領域に存在する。先にも述べた通り、枯草菌のRecA配列を確認すると、MuBのLoop 1の相同領域付近において、RecAでは16 aaの挿入がある。つまり、RecAの方がloop L1の後ろに続く長さが長いと言える。この長さの違いが、MuBには見られない、RecAのssDNA結合/dsDNA結合の二つのモードを生み出す要因になっているのかもしれない。

また、*muB*が紫外線耐性を宿主に付与しないことは (図11)、系統解析から近縁であることが示された*uvsX*と比較すると大変興味深い。*uvsX*は、バクテリオファージT4が持つdNTPaseをコードする遺伝子である。バクテリオファージT4も、Muと同じMyoviridae科に属す、二本鎖DNAウイルスである。興味深いことに、*uvsX*は*recA*のように、相同なDNA鎖の対合を行うことが知られており[73]、この反応には*uvsX*を含むT4ファージにコードされたいくつかの遺伝子産物が関与している[74]。特筆すべきは、*uvsX*遺伝子が、*recA*欠損の大腸菌に対して、野生株よりも高い紫外線耐性を与えるという知見である[75]。この細胞では、 $\lambda$ ファージと大腸菌ゲノムの組換え頻度が上昇するが、溶原化している $\lambda$ ファージの誘導は抑制されている[75]。筆者の行った実験において、*muB*と*recA*を置換した枯草菌の紫外線耐性は、*recA*欠損の枯草菌と同程度であり (図12)、紫外線耐性の付与は認められなかったという点は、*uvsX*とは異なるのである。*uvsX*が宿主に紫外線耐性を与えるのに対し[73]、*muB*では

その効果が実験的に認められなかったこと、及び*muB*と*UvsX*が分子系統学的に近縁であることを考慮すると、*muB*の方がより祖先配列に近い配列を保存している可能性が考えられる。UvsXのドメイン構造はRecAやMuBと同様、3つのドメインからなり、UvsXにおけるRecAと相同な二つのssDNA結合領域 (loop L1/L2) が同定されている[76,77]。MuBとUvsX (NP\_049656.2) をEMBOSS Stretcherを用いて比較したところ、互いの相同性は Identity: 14.1%, Similarity: 31.2%となった (図13)。それぞれある3つのドメインは枯草菌のRecAと比較した場合と同様、互いに近傍に整列したが、MuBのLoop 1は枯草菌のRecAのloop L1と重なるように整列されたのに対し、UvsXのloop L2と重なるように整列された。枯草菌RecAとMuBの比較結果も考慮すると、MuBに紫外線耐性付与能が見られない原因は (図11)、RecAやUvsXに2つ見られるLoop構造がMuBには1つしか無いためであると考えられる。また、MuBのLoop 1配列が枯草菌RecAのloop L1周辺、及びUvsXのloop L2周辺と相同であることから、第3章 第1節に示した分子系統解析結果を考慮すると (図8)、*recA*遺伝子及び*UvsX*遺伝子上のloop L1 / L2構造をコードする配列は、*muB*遺伝子に近い祖先配列のLoop 1構造をコードする配列から互いに独立に進化したことを示唆すると考えられる。

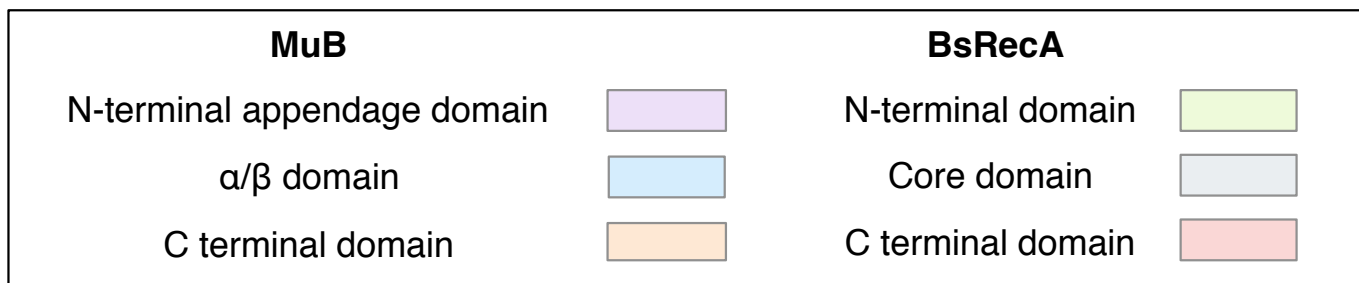
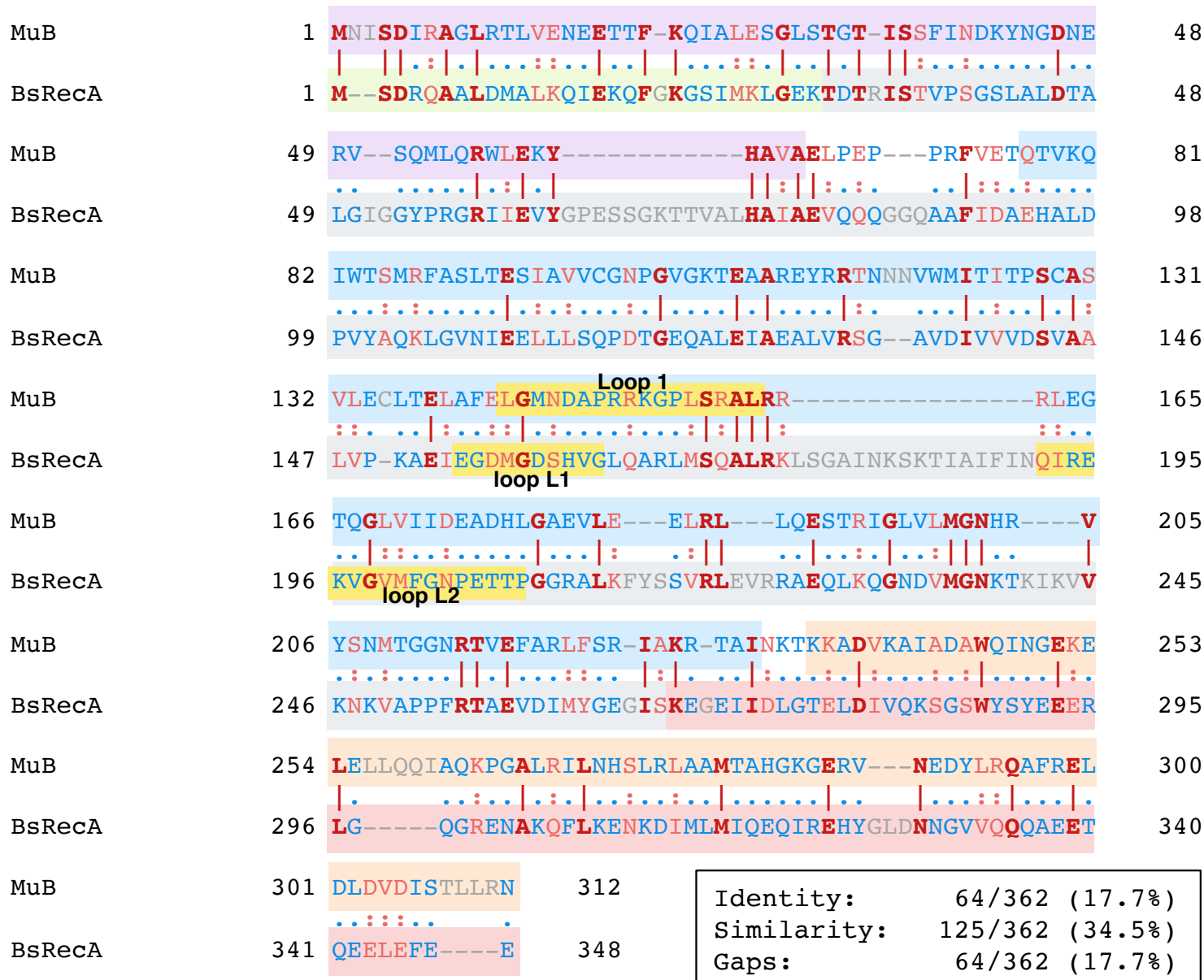


図12. MuBと枯草菌RecAのアミノ酸配列比較解析

EMBOSS stretcherによるMuB (NP\_050608.1) と枯草菌RecA (BsRecA, WP\_003245789.1) のアライメント結果を示す。MuBとBsRecAでアミノ酸残基が一致したもの及び類似のものは赤系の文字, その他の文字を青で示す。ただし, Gap領域のアミノ酸残基はグレーで示した。黄色でハイライトされた配列は, MuBまたはBsRecAのDNA結合モチーフを示す (MuB: Loop 1, BsRecA: loop L1 / L2)。アライメント右下端のボックスは, アライメント結果に基づく相同性を数値化したもの (Identity: アミノ酸が一致した数及び割合, Similarity: アミノ酸の性質が一致した数及び割合, Gaps: アライメントされなかったアミノ酸の数及び割合)。アライメントの背景色はMuB及びBsRecAのドメインを示し, 下端のボックスと対応している。

MuB	1	MNISD--IRAGLRTLVE-----NEETTFK-----QIALESGLSTG	33
		:..:  : ... ..   :..:  .. ...:..	
UvsX	1	MSDLKSRLIKASTSKLTAELTASKFFNEKDVVRTKIPMMNIALSGEITGG	50
MuB	34	TISSFI-----NDKYNGDNERVSQMLQRWLEK-----YHAVAELP----	68
		.. ..:  .. ... .. ...:..:  .. ...:..	
UvsX	51	MQSGLLILAGPSKSFKSNFGLTMVSSYMRQYPDAVCLFYDSEFGITPAYL	100
MuB	69	-----EPPRFVET--QTVKQIWTSM--RFASLTESIADVCGNPGVVKTEA	109
		: .:.:  : : :..:  ..:..:..:  .. ...:..	
UvsX	101	RSMGVDPERVIHTPVQSLEQLRIDMVNQLDAIERGEKVVVFIDSLGNLAS	150
MuB	110	AREYRRTNN--VWMITITPSCASVLECLTELAFE-----LGMNDAPRRK	152
		.. ...:  ..: ...:..: ...: ...:  .. ...:..	
UvsX	151	KKETEDALNEKVVSDMTRAKTMKSLFRIVTPYFSTKNIPCIAINHTYETQ	200
		loop L1	
MuB	153	GPLSRALRRRLEG--TQGLVVI-----DEADHLGAEVL--EELRLL	189
		.. :..:..:  ..:..:  .. ...:..:  .. ...:..	
UvsX	201	EMFSKTVMGGGTGPMYSADTVFIIGKRQIKDGSDLQGYQFVLNV EKSRTV	250
		loop L2	
MuB	190	QESTR--IGLVLMGNHRVYSNMTGGNRTVEFARLFSRIAKRTAINKTKKA	237
		: :..:  :..:..:  .. ...:..:  .. ...:..:  .. ...:..	
UvsX	251	KEKSKFFIDVKFDGGIDPYSGLLDMALELGFVVKPKNGWYAREFLDEETG	300
MuB	238	DVKAIADAWQINGEKELELLQQIAQKPGALRILNHSRLAAMTAHGKER	287
		: ...:  :..:..:..:  ..:..:..:  .. ...:..:  .. ...:..	
UvsX	301	EMIREEKSWRAKDTNCTTFWGPLFKHQPFDAIKRAYQLGAIDSNEIVEA	350
MuB	288	VNEDYLRQAFREL-----DLDVDISTL-----LRN	312
		..:..:..:..:  ..: :..:  .. ...:..	
UvsX	351	EVDELINSKVEKFKSPESKSKSAADLETDLLEQLSDMEEFNE	391

Identity:	64/362 (14.1%)
Similarity:	125/362 (31.2%)
Gaps:	64/362 (20.2%)

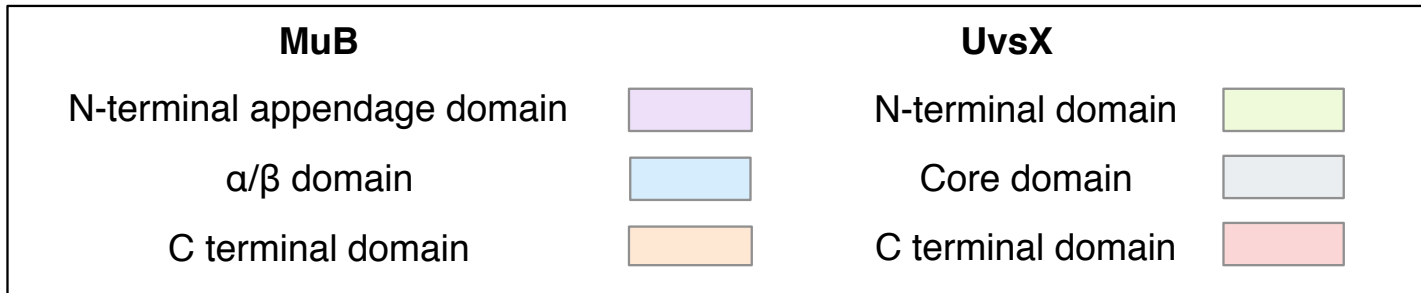


図13. MuBとUvsXのアミノ酸配列比較解析

EMBOSS stretcherによるMuB (NP\_050608.1) とUvsX (NP\_049656.2) のアライメント結果を示す。表記方法は図12に準ずる。ただし、黄色でハイライトされた配列は、MuBまたはUvsXのDNA結合モチーフを示す (MuB: Loop 1, UvsX: loop L1 / L2)。

# 第4章 本論-3

ゲノム上の突然変異, そのランダム性について.

# 突然変異とランダム性

これまで第2章と第3章で述べた研究結果から、IS256Bsu1にコードされたトランスポゼーンによる挿入配列の転移には、相同組換え酵素遺伝子*recA*が必要であることが分かった。ただし、*recA*がこれまでよく研究されてきた相同組換えは、この転移反応とは関係が無いようである。さらに、バクテリオファージMuにコードされた*muB*は*recA*の本実験における転移反応への必須性を相補することが分かった。相同組換え遺伝子と*muB*を含む網羅的な分子抵当解析の結果から、*muB*が既存の*recA*遺伝子から派生したという可能性は低く、バクテリオファージ Muとその近縁のウイルスが保存している*muB*ホモログは一つのクレードを形成し、その近傍に存在するクレードがバクテリオファージT4に代表されるウイルス群の持つ*uvsX*であった。*muB*と*recA*を置換した株への紫外線照射実験から、*muB*には、*uvsX*や他の相同組換え遺伝子のような、宿主に紫外線耐性を付与する能力は見られなかった。他方、未だ疑問が残る。系統解析結果を考慮すると、相同組換え酵素遺伝子は出鱈目に変異を繰り返しながら都合の良いものが残って来たとは考えにくい。第一、もしランダムに変異が生じても遺伝子が宿主にとって有用で在りさえすれば良いのならば、なぜ*recA*や*radA*の分岐に従ったGC含量の増加が見られるのだろうか。DNA上の突然変異がランダムに生じるならば、4種類の塩基の置換は等確率に生じるはずである。そこで本章では、より発展的な研究を交えつつ、ゲノム上の突然変異の生ずる様式について考えてゆく。

## 1-1. 突然変異とランダム性が結びつくことになる歴史的背景

まず、突然変異とそのランダム性が関連づけられた経緯について振り返る。突然変異がランダムに生じる可能性を示したのは、Salvador Edward LuriaとMax Ludwig Henning Delbrückである。両博士は1942年、大腸菌のウイルス抵抗性突然変異体の出現と遺伝性を継代培養により確認し、ウイルスに抵抗性の大腸菌の現れ方がランダムであり、さらにこの形質が遺伝することを示した[78]。注意したいのは、ここでの「ランダムである」対象は「どの大腸菌か」ということ、そしてその実験が行われたのが1942年、つまり生物の遺伝情報の



担い手がDNAであることが 1944 年Oswald Theodore Avery, 1952年にAlfred Day HersheyとMartha Cowles Chaseによって其々示され, 1953年, その立体構造がRosalind Elsie Franklinのデータに基づきJames Dewey WatsonとFrancis Harry Compton Crickらによって提案されるよりも前であるということである[79-81].

ここで言う突然変異はいわゆる自然突然変異 (spontaneous point mutations) の事であるが, その原因は, 現在では塩基が稀に互変異体を形成することにあると考えられている[82,83]. 核内互変異性の場合, 環構造を持つ分子のプロトン転移により異性体を生じるが, これによりDNAでは通常見られないAC対とGT対を発生する要因になるとされる. ここから, 核酸分子が等しく一定確率で互変異体を生じるならば, ゲノム上における自然突然変異の発生がランダムに起こると考えられるようになったと推察される. しかしながら, 現時点では, 細胞内において核酸の互変異体がどの程度生じているのか, それが突然変異に寄与する割合はどの位なのかは不明であり, 今の所この言説は否定も肯定もされないものと思われる. 互変異体と自然突然変異を結びつける仮説の問題点は過去に言及があり, (a) 生理的なpHに起因する「稀な」互変異体の形成という仮定に証拠がないこと, (b) 自然突然変異はトランジション変異だけでなくトランスバージョン変異も存在するが, これは互変異体の形成だけでは説明できないことが挙げられている[82]. 近年の, 互変異体の形成を医療に応用した例にKP1212が挙げられる. HIV薬として開発されたKP1212はDNAのアナログであり, pHが7より大きい条件において5つの異なる互変異性体となることを利用してHIVの突然変異頻度を上昇させ, ウイルスを不活化に追い込む[84]. ただし, KP1212自体はあくまでも核酸のアナログであり, pHが互変異性体の変換率を高める事実が細胞内のDNAにおいてどの程度適用可能かどうかは依然として不明なままである. Danaらは, 大腸菌のDNA Polymerase I (Klenow fragment) のE710A変異体が*in vitro*においてA-dCTP, G-dTTP (鋳型は一文字, 対合する塩基はdNTPとして表記) の誤対合が野生株よりも顕著に高いことを示し, その原因は活性部位の形状が変化することで, 取り込んだ塩基を正確に配置できないためである可能性があることを, 好熱菌の立体構造を参照しながら議論している[85]. この時, 他の組み合わせの誤対合が生じにくいのは活性ポケット内での塩基と酵素のフィットの具合に依存しているとしているが[85], これは取りもなおさず, 突然変異が単なる間違いなどではなく, DNA複製酵素による反応産物の一つであることを意味している.

## 1-2. ランダムな定義

ここで「ランダム」という言葉の定義を見直す。ランダムである、という言葉が一般に使われる際、人為性のなさの他に、「一見してよく分からない」ということにまでその意味が拡張されることがあるが、この言葉を数量的な議論の下で扱う場合、その意味は観察対象の現れる確率が互いにおよそ均等であることを意味する。ある遺伝型のバクテリアが薬剤に抵抗性を示す時、その出現する数が単位あたり常に一定ならば、「どの細胞が変異するかはランダムな事象である」ということになる。では、ゲノム上の突然変異は本当にランダムに生じているのだろうか。

# GC含量と高頻度突然変異株の研究

GC含量は、生物の持つゲノムの性質を理解する上で役立つ指標の一つである。現在では、バクテリアのゲノム情報（数Mbpほど）であれば、比較的迅速に決定することが可能であり、GC含量とその生物の持つ遺伝子の特徴などを関連づけて研究することが可能である。ここでは、筆者が行なったゲノムのGC含量とDNA修復因子欠損株における突然変異の偏向性に関する研究について述べる[86].

## 2-1. バクテリアにおけるDNA修復, 複製, 校正因子の保存性の検証

### <方法, 結果>

筆者は、突然変異に関わる遺伝子について探索、データベース解析を行う過程で、放線菌群（高GCグラム陽性菌）がDNA誤対合修復因子をコードする遺伝子*mutS*を欠損していることに気づいたが、この発見をきっかけとし、MutS, MutL, MutTという三種類のDNA修復関連因子に加え、DNA複製因子PolCまたはDnaE, DNA校正因子DnaQの保存性の検証を、70種類のバクテリアに対して行った。方法は次の通りである。大腸菌(*Escherichia coli* str. K-12 substr. MG1655, Taxonomy ID: 511145)に保存されたこれらの因子（DnaE: NP\_414726, DnaQ: NP\_414751, MutS: NP\_417213, MutL: NP\_418591, MutT: NP\_414641）をクエリとし、70種類のバクテリアにおけるゲノム情報に対して、DELTA-BLAST [87]による網羅的な情報検索を実施した (<https://www.ncbi.nlm.nih.gov>)。その結果、バクテリアにおけるこれらの遺伝子の保存性に、次の特徴が見られることが分かった（図14）。

- 1) データベース上のすべての放線菌に属すバクテリアはMutSとMutLを持たない（図14）。
- 2) フェルミクテス門のバクテリア（低GCグラム陽性菌）は他のバクテリアとは異なり、DNA複製ドメインと校正ドメインを持つPolCをDnaE, DnaQとともに保存する（図14）。

## 2-2. 枯草菌*mutS*欠損株が示す*rpoB*遺伝子上の突然変異の傾向.

放線菌がDNA修復因子MutS, MutLを持たないことと, 放線菌のゲノムGC含量が高いことが関係しているならば, 低GC含量のゲノムを持つバクテリアでこれらの因子をコードする遺伝子が欠損した場合, ゲノム上の突然変異にはどのような変化が見られるだろうか. これを確認するため, 次の実験を行なった.

### <方法, 結果, 考察>

低GCグラム陽性菌の一種である枯草菌168株において*mutS*遺伝子を欠損させた株を作製し, 野生株と*mutS*欠損株の間で抗生物質リファンピシン抵抗性の株を選択, リファンピシン耐性変異が多く観察される*rpoB*遺伝子上の突然変異をシーケンス解析により調べた (野生株: 20株, *mutS*: 30株). その結果, 野生株における*rpoB*上の突然変異はG/CからA/Tへの変異が多いのに対し (C-A:11%, C-T:56%, A-G:33%), *mutS*を欠損した株の場合はA/TからG/Cに偏っていた (C-T:14%, A-G:86%) [86]. *mutS*欠損株の場合, リファンピシン耐性菌の出現頻度は $10^{-6}$ であり, これは, 野生株におけるその割合 ( $10^{-8}$ ) と比べて100倍高い [86]. 従って, *mutS*欠損株がバクテリア集団に存在する場合, この実験条件ならばG/Cに偏った突然変異を集団中に生じやすいと考えられる. 言い換えれば, DNA修復因子を欠損した場合はGCに偏った変異が生ずるということである. 枯草菌の野生株における突然変異がAやTに偏ることで, 低GC含量のゲノムが維持されているのかも知れない. これらの実験結果を, ゲノムのGC含量にまで拡張するのは現時点では容易ではないのだが, DNA修復因子の保存性がゲノムのGC含量に影響している可能性はより高まったと考えられる.

では, この実験からゲノム上の変異とランダム性の関係についてはどのようなことが考えられるだろうか. 突然変異の生じた座位を確認すると, 野生株は6種類あったのに対し, *mutS*欠損株は3種類であり, 互いに有意に異なる値を示す ( $p < 0.01$ ,  $\chi^2$  test, 表3). そのほとんどが*rpoB* A1406G (46.7%) と*rpoB* A1445G (36.7%) に集中し, これらの値は野生株よりも有意に高い値であるが, 反対に他の4つの座位の値は野生株の方が*mutS*欠損株よりも有意に高い値を示す (いずれも $p < 0.05$ , 残差分析, 表3). これらの結果は, 突然変異のゲノム上におけるランダム性を明らかに否定している. 何故ならば, リファンピシン耐性菌の出現がランダムであり, かつ*mutS*欠損により変異の修復が行われなければ, 得られる耐性菌の遺伝型が特定の座位に偏ることはないからである. *mutS*欠損でよく観察されたリファ

ンピシン耐性菌と同じ*rpoB*変異を持つ株は野生株由来のサンプルからも観察されているため (*rpoB* A1406G: 20.0%, *rpoB* A1445G: 5%) , 培養液中でこれらの変異を持つことが増殖に有利であるとは考え難い. 以上をまとめると, 少なくとも本研究の結果は, DNAのゲノム上の変異はランダムではなく, DNA修復因子に関わる遺伝子の遺伝型によって傾向が変わるものであることが示唆された. ゲノムのGC含量の決定要因は結局のところ定かでは無いが, 例えば, *mutS*や*mutL*のようなDNA修復因子の組み合わせにより決定されると考えられる.

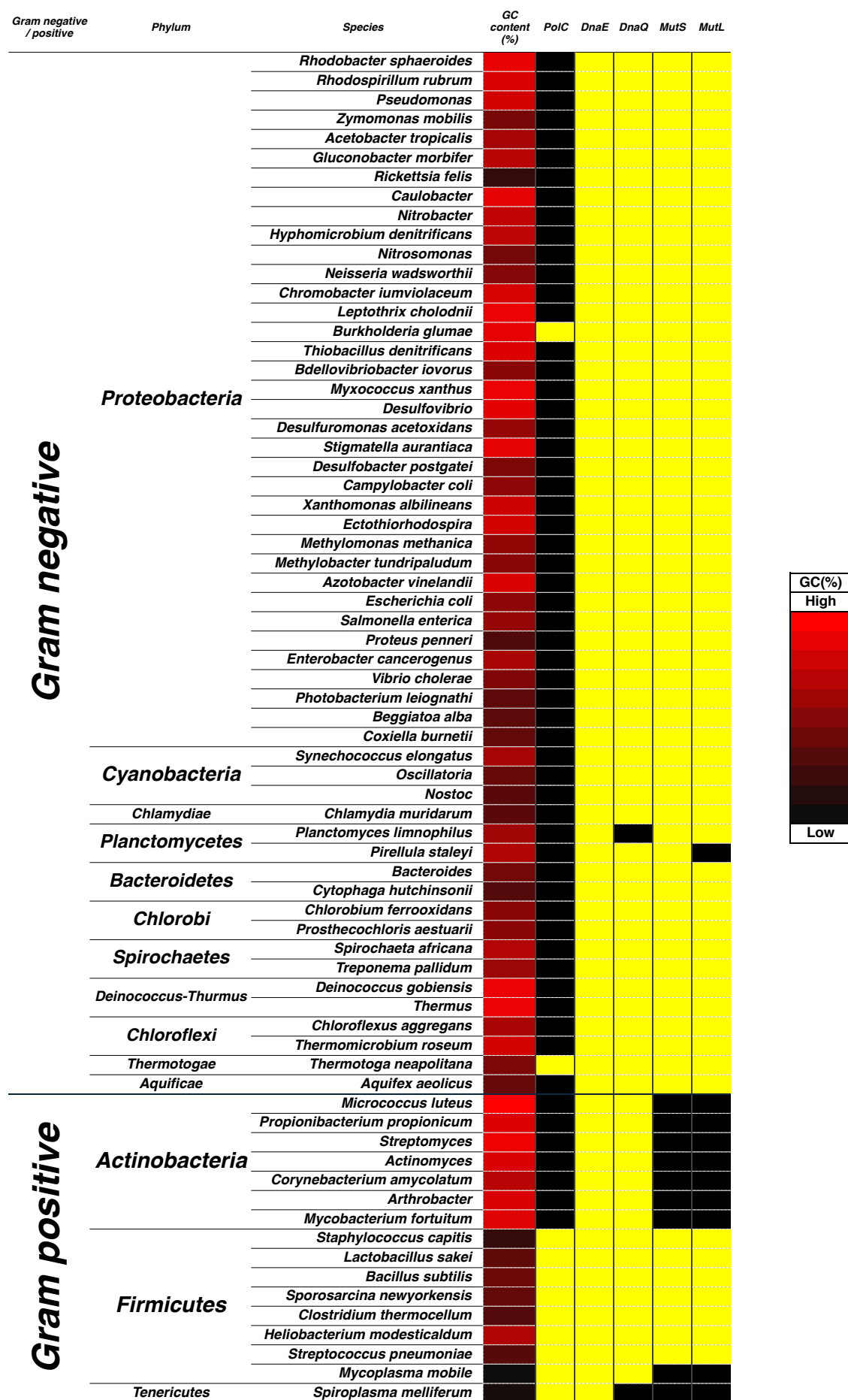


図14. バクテリアにおけるDNA修復, 複製, 校正因子の保存性

表に示す種におけるゲノムのGC含量 (GC content) は、左端のカラースケールと対応しており、GC含量が高いものほど赤く表示されている。右側5列はDNA複製因子 (PoIC, DnaE), DNA複製校正因子 (DnaQ), DNA修復因子 (MutS, MutL) の保存性を示す。各生物種でこれらの因子が保存されている場合、セルの色は黄色となっている。

<b>Rif<sup>r</sup> Mutation</b>	<b>Frequency (%)</b>	
	<b>WT</b>	<b><i>mutS</i></b>
<i>rpoB</i> A1406G	20.0	46.7
<i>rpoB</i> A1445G	5.0	36.7
<i>rpoB</i> C1433A	10.0	0.0
<i>rpoB</i> C1444T	25.0	13.3
<i>rpoB</i> C1460T	25.0	0.0
<i>rpoB</i> G1454A	5.0	0.0
No mutation	10.0	3.3

**表3. 枯草菌野生株と*mutS*欠損株における各種Rif<sup>r</sup>菌の出現頻度の違い**

枯草菌野生株 (WT) と*mutS*欠損株 (*mutS*) における, 遺伝型の異なるRif<sup>r</sup>菌の出現頻度 (%) を示す. 「Rif<sup>r</sup> mutation」は各Rif<sup>r</sup>菌の*rpoB*遺伝子上の変異を示す. ただし, 「No mutation」は, シークエンスを確認した*rpoB*遺伝子上に変異が見られなかったものである.

## 第4章 第3節

# 突然変異が特定のゲノム構造内に存在する高頻度突然変異領域 (Hotspot) に発生したと仮定した場合におけるゲノム上の突然変異分布状況のシミュレーション

### 3-1. 突然変異が特定のゲノム構造内に存在する高頻度突然変異領域 (Hotspot) に発生したと仮定した場合におけるゲノム上の突然変異分布状況のシミュレーション

近年、合田徳夫博士は、志波優博士により精査されたヒトゲノム上に見出されている全SNPsについて、その座標間距離の頻度を調査することで、SNPsの分布がランダムではなく、べき分布であることを発見した。更に、ボックスカウント法（フラクタル構造を解析する手法の一つ）を用いて、ゲノム上のSNPsの分布が解像度によらず一定である、つまりフラクタル性を有することを明らかとした。これらの事実は、1)あるSNPsの近傍には別のSNPsが存在する可能性が高い、2)SNPsのゲノム上の分布には、どのようなスケールにおいても共通して確認可能な特定の分布構造（または分布様式）があることを意味する。

ここで、簡単にフラクタルについて紹介する。自然界には、様々な幾何学的構造を持つ物質が存在するが、中にはその物質に特徴的な構造を更に拡大して観察した場合、元の倍率で観察した時と同じ構造が見出されることがあり、このような性質を自己相似性という。フラクタルは、このような自己相似性を有する構造を表す概念であり、マンデルブロ (B.B.Mandelbrot) により提唱された。自然界におけるフラクタル構造は、例えば河川網や稲妻、雪の結晶、ロマネスコ・ブロッコリーなどに見られる[88]。

さて、合田博士により解析されたSNPs (HapMap, iJGVD; Integrative Japanese Genome Variation Database, 1000 Genomes) は、当然ある一個人のゲノムについて見ていくわけではないので、SNPsのゲノム上の分布に偏りがあるということは、SNPsが高頻度に観察されるゲノム領域が、ヒトのゲノムに共通した高頻度変異領域 (Hotspot) であることを意味する。また、1000 Genomesや iJGVDのデータに関する報告によると、同義置換変異 (Synonymous SNPs) と非同義置換変異 (Non-Synonymous SNPs) の割合はおよそ一対

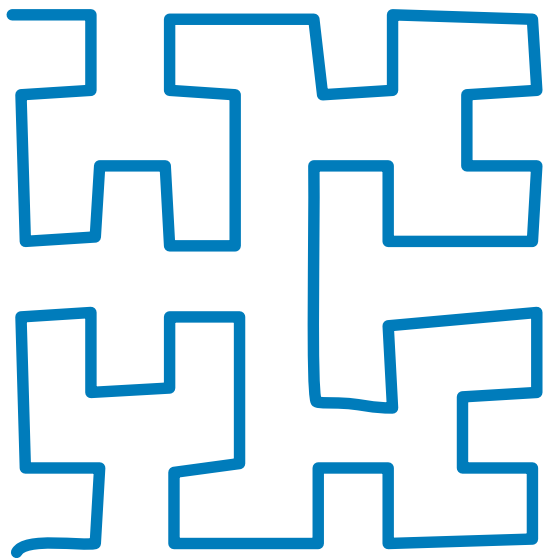


一の割合であり[89,90], SNPsの分布の偏りが自然選択によるものではないことを示唆している。では, ヒトゲノムのSNPsに見られる分布の偏り, およびその繰り返し構造は何に由来するのだろうか。興味深いことに, 近年のヒトゲノムの物理的構造の研究により, 核内のゲノムの三次元構造は, ヒルベルト曲線と呼ばれるフラクタル構造の一種に準ずるものであることが明らかになった[91-93]。ヒルベルト曲線は, 線が一切交差することなく平面や空間を充填可能という幾何学的な特徴を持つフラクタル構造の一種である[91]。

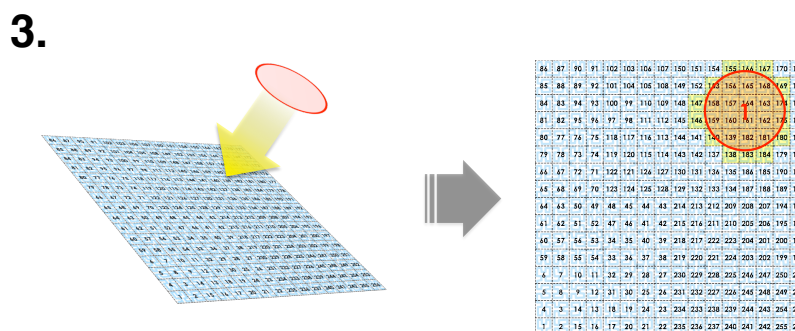
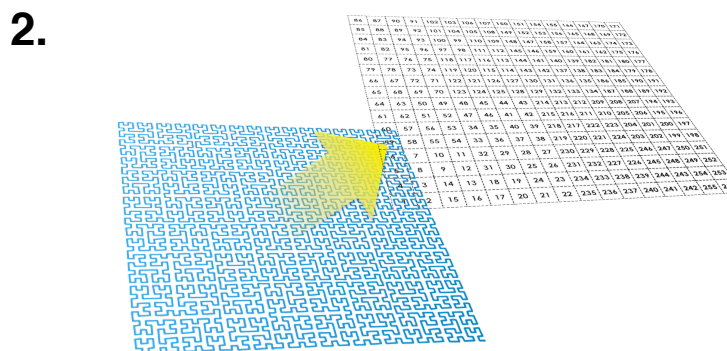
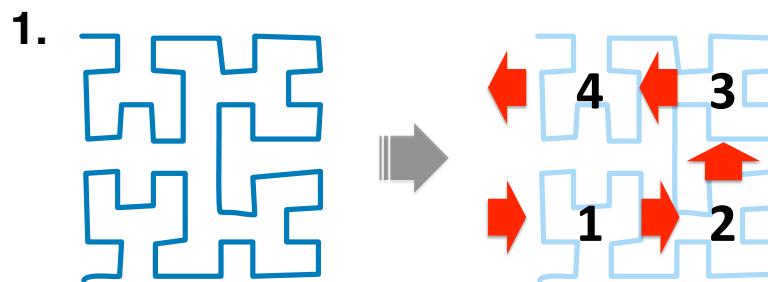
### <方法, 結果, 考察>

筆者は, この報告および仮定された高頻度変異領域の存在に基づき, 二次元のヒルベルト曲線上に高頻度変異領域に相当する閉じた線(円など)を配置し, その内部および交点の座標が突然変異の生じた座位であるとした場合, フラクタル性が見出され, かつその分布がべき分布となるか, シミュレーションを行った(図15)。シミュレーション方法および詳細な結果は付録を参照のこと。その結果, 円形によりヒルベルト曲線を切り取った場合, フラクタル構造の有無をボックスカウント法により確認すると, 円の数や円の大きさを変更してもフラクタル性が見出された(図16. A, B)。また, 座標をランダムに選択して同様の解析を行った場合, 得られた近似直線の傾きは円の場合よりも小さく, 相関係数は著しく低下した(付録8を参照)。これらの解析は, 二次元から三次元へと次数をあげた場合, 円形は球形となるが, この場合でもやはり結果は同じとなると予想される。以上をまとめると, ヒトゲノムにおけるSNPsのゲノム上での分布はランダム分布ではなくべき分布であり, その分布にはフラクタル性があることが分かった。ヒトゲノムが細胞内でヒルベルト曲線に似た三次元構造を取る時, ローカルな高頻度変異領域がその内部に存在すると仮定すると, ヒトSNPsの分布と同様の観測結果が得られる。従って, ヒトゲノムには, その物理的構造に基づく高頻度変異領域が存在することが示唆された[94]。

**A**



**B**



**図15. 突然変異分布状況のシミュレーション**

- A. ヒルベルト曲線及び、シミュレーションの際の数え上げの単位となる構造を示す。
- B. シミュレーションのマトリクス作製及びシミュレーションの手順を示す. 1) マトリクスの各セルへの番号付のルール. 2) 16x16のマトリクスとなるようにヒルベルト曲線を作製し、マトリクスを重ねて番号を振った. 3) 作製したマトリクスに任意の大きさの円を乗せ、円に接するか内包されたセルの番号を記録した.

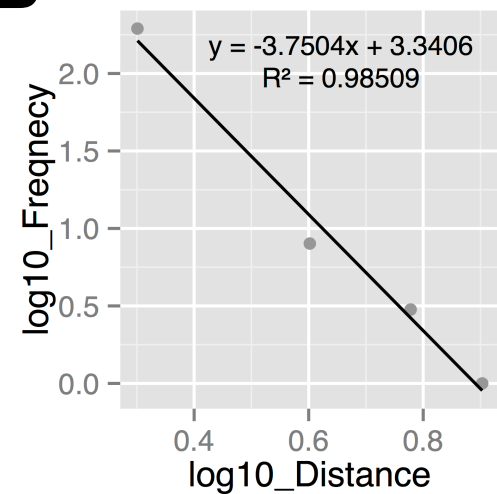
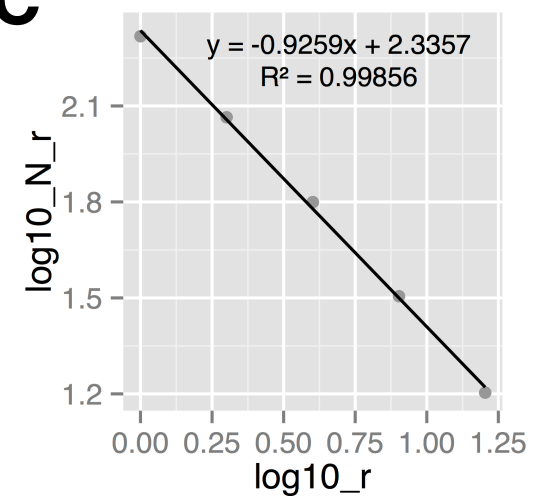
**A****B****C**

図16. 突然変異分布状況のシミュレーション

A. シミュレーション解析に用いた図形の例[94].

B. A図のシミュレーションにより得られた座標間距離の頻度を求め、頻度と距離について対数をとってプロットしたグラフ (log10\_Frequency: log10(頻度), log10\_Distance: log10(座標間距離)) . 上端の数式はデータに対する線型近似直線, R<sup>2</sup>は相関係数.

C. A図のシミュレーションにより得られた座標の数をボックスカウント法で求め、解像度ごとにプロットした図 (log10\_N\_r: log10(各解像度における座標の数), log10\_r: log10(解像度)) . 上端の数式はデータに対する線型近似直線, R<sup>2</sup>は相関係数. グラフの傾きの絶対値 (0.9259) はフラクタル次元を意味する.

### 3-2. 小括

以上のシミュレーションにより、ヒトゲノム上のSNPsの分布がべき分布となる原因が、細胞内における立体的なゲノムの構造内に、高頻度変異領域（Hotspot）が局在することによるということが示唆された。これらの結果は、突然変異がゲノム上でランダムに生じないことを意味する。ゲノム上の突然変異がランダムでない可能性に言及した研究は他にも存在する。大腸菌34株のゲノムを比較し、12万以上のSNPsを比較した研究では、ゲノム上にSNPsが多い領域（hot region）と少ない領域（cold region）が存在することが報告されている[95]。特に興味深いのは、同義置換変異の多様性の高さを数値化して比較した場合、多様なSNPsを有する遺伝子領域は互いに近隣に存在することである[95]。このような領域の遺伝子を精査すると、ゲノムDNAにおける遺伝子のコードされている向きが一致する場合が多く、その遺伝子機能に類似性が見られる[95]。さらに、遺伝子の転写レベルとSNPsの多様性を比較すると、転写が頻繁に行われる遺伝子ではSNPsの検出数とその多様性は低下している[95]。合田徳夫博士により見出されたヒトゲノム上のSNPsの分布の偏りの存在は、その遺伝子発現との相関性までは確認されていないが、SNPsがゲノム上に不均一に現れるという点で、大腸菌からヒトまで共通していることを示唆している。

## 第4章の考察と総括

以上2つの突然変異に関する研究から、ゲノム上の突然変異、選択という現象はランダムに生じているわけではないことが示唆された。これらは、第3章において生じた疑問、即ち、相同組換え酵素遺伝子の系統とGC含量に相関性が見られる要因に、どのような示唆を与えられるだろうか。

相同組換え酵素遺伝子のGC含量と系統関係について調査した結果、一部のバクテリアの *recA* と古細菌の *radA* のGC含量が系統関係の相対的な距離に比例し、かつ宿主のゲノムGC含量に依存していたが（図9）、その要因は何であろうか。DNAの結合におけるAT対やRNAの結合におけるAU対は、それらのGC対と比較して、水素結合の数が少ない。そのため、温度の上昇に伴い、AT対の多い二本鎖DNAなどは、少ないものと比べて解離しやすい。しかし、ゲノムDNAのGC含量と生息環境の温度の間に相関性は見られないことが知られている[96,97]。一方で、古細菌において、tRNAのGC含量が生育温度と高い相関性を示すが、tRNAとゲノムのGC含量には相関性がないことが報告されている[97]。このことは、結果として、古細菌の細胞内でのタンパク質合成反応が、細胞の生息する環境に強く依存することを示唆している。従って、第3章で示した相同組換え酵素遺伝子の系統間におけるGC含量の違いは（図9）、宿主のtRNAに影響されたものではないと考えられる。バクテリアについては、宿主に固有の変異傾向があることから[86]、各細胞が相同組換え遺伝子を獲得した後、一定のGC含量に収斂したと考えられる。この時獲得した相同組換え遺伝子は、それぞれの宿主のゲノムのGC含量に近いものであったかも知れない。

突然変異、選択に傾向が見られることは、実験室内の研究に留まらない。近年報告されたタスマニアデビルの突然変異と進化に関わる研究では、タスマニアデビルの集団に広まってしまった感染性の腫瘍（DFTD; Devil facial tumour disease）に抵抗性の変異を、たった20年、世代にして4から6世代の間に獲得しつつあることを報じている[98]。タスマニア島に存在する三つのタスマニアデビルの群からそれぞれサンプルを採取し、DFTDに抵抗性を示すためのゲノム領域2ヶ所を特定し、その領域の遺伝子7つの内5つが発ガンもしくは免疫

システムに關与することを見出した。興味深いのは、この変異が特定の群れで生じ、交配によって3つの群れに広まったのではなく、3つの群れにおいて同時多発的にこの変異が獲得されたということである。このことは、DFTDの形質に由来する細胞内の環境が、この突然変異の発生の由来であることを示唆している。筆者の行った突然変異のGC含量偏向性とミューテーターの研究では、ミューテーターとなる遺伝型 (*mutS*) を持つ枯草菌168株の *rpoB* 遺伝子上の突然変異の傾向が、ATではなくGCに偏ることのみならず、特定の座位における突然変異選択率が格段に上昇することを示唆した[86]。また、相同組換え因子の系統関係と、そのGC含量の相関性の解析結果は、各系統の分岐回数の増加に伴って、GC含量が多くなることを示した (図9)。このことは、細胞の生息環境のみならず、細胞の内的環境が染色体の突然変異の傾向を生む要因であることを意味すると考えられる。

ゲノム上の変異がランダムに生じるわけでは無いと言うと、木村資生博士、太田朋子博士の中立説 (ほぼ中立説) と異なる立場であると誤解を受けることがある[99-101]。しかし、木村博士は、ゲノム上の変異がランダムに生じる、とは述べていない。ここでは、中立説とゲノム上の変異の分布のランダム性の関係性を否定する上で重要な要点を挙げる。まず、第一に、中立進化説は、一つの遺伝子を複数の生物で比較した時、その推定分岐年代から逆算すると、時間あたりの変異回数が一定であり、しかもその変異の多くがアミノ酸配列上の変異を伴わない、というものである。中立であるとした理由は、アミノ酸配列上の変異を伴わないならば、表現型に影響しないから、自然選択を受けないためである。第二に木村博士は、遺伝子ごとに、時間あたりの変異蓄積数が異なることを示している。これは取りも直さず、ある個体の細胞内にある染色体において、その座位ごとに変異の発生確率が異なること、つまり、あるゲノム上に発生する変異がランダムでは無い (座位ごとに等確率でない) ということを意味する。非選択的な変異の、時間あたりの蓄積率が異なるならば、ゲノム上の変異はランダムに生じているとは言えない。第三に、自然選択との関係であるが、アミノ酸の活性に関わる遺伝子上のコドンでは変異が発生し辛く、これは変異が起きても選択により淘汰されるためであると考えられている。筆者のGC含量に関わる研究は、抗生物質による選択がかかった条件でもなお、細胞内の選択に関わる座位 (*rpoB* 遺伝子の region 1) に見られる変異は、その細胞の遺伝型によって異なっており (表3)、このことは特定の遺伝型が変異の種類や生じる座位の決定に影響する可能性があることを示唆している[86]。

繰り返しになるが、ランダムであるとは、観察される事象の発生確率が互いに等しいという意味である。リファンピシンを用いた枯草菌野生株と*mutS*欠損株における薬剤耐性菌出現頻度が示すように、純粋な遺伝型を持つ集団において、つまり実験などの状況のみにおいて、どの細胞が突然変異体となるかはランダムな事象であり、各遺伝型に固有の確率が存在するのかもしれない。一方、リファンピシン耐性変異のパターンと遺伝型の対応関係、およびヒトSNPsの分布の研究より、ゲノム上の各座位における変異確率が等しくはなく、かつその確率の偏向にはゲノムの物理的な構造および宿主の遺伝型が影響していると考えざるを得ない。

以上、3章からなる本論により、相同組換え酵素遺伝子がトランスポゾンの転移に関与する遺伝子であり、しかもバクテリオファージにコードされるDNA結合タンパク質遺伝子と共通祖先を有することが示唆された。相同組換え酵素遺伝子は通常の遺伝とは異なる様式、つまり水平伝播または感染を介して細胞間を移動し、宿主のゲノムに近いGC含量を持つに至ったと考えられる。ゲノム上の突然変異には具体的な様式が存在せず、出鱈目に変異し、適応的なものが偶然残るといった考え方が未だ根強いが、少なくとも筆者らの研究では、ゲノム上の突然変異には固有の偏りが存在することが示唆された。では、これらの実験事実が示す、進化学的意義とは一体何か。次章では、これらの実験による結果を背景に、分子系統解析結果に基づく相同組換え遺伝子の進化の経緯を紐解いてゆく。

# 第5章 総合討論、及び結論

相同組換え酵素の進化プロセスについて.



# 総合討論

## 相同組換え遺伝子を細胞が獲得する進化プロセスの構築

本研究の目的は、生物が相同組換え遺伝子を獲得したプロセスを構築することにある。相同組換え反応は、DNAの修復に加え、ある生物が自身のゲノムDNAに新たな遺伝情報を加える事を可能としている。新たな遺伝情報の獲得は、生物のゲノム配列が変化するよりもはるかに効率的に、新規の機能を獲得する手段となる。では、そもそも、生物はその相同組換え遺伝子をどのようにして手に入れるに至ったのか。本節では、これまでの研究結果をもとに、相同組換え遺伝子の成り立ちについて、実際に辿った経路とその獲得方法という二つの視点から考察する。

### 1. 相同組換え遺伝子の祖先配列を有した宿主について

まず、RecAとMuBに共通点に基づき、相同組換え遺伝子の祖先配列の由来について考える。MuBはEnterobacteriophage Muのゲノム上の*muB*遺伝子にコードされたdsDNA結合タンパク質であり、Mutator like トランスポゼースであるMuAとMu phage genomeの複合体を新たなゲノム領域へ呼び込む働きを担う[19,58]。一方、RecAとDNAの結合様式は二種類あり、RecA-ssDNA複合体の形でdsDNAに結合して相同鎖の探索を行うだけでなく、dsDNAに直接結合する場合が存在する[61]。いずれの状態を取るかはRecA loop L1 regionにあるアスパラギン酸（大腸菌ではD161）により決定される [50]。つまり、RecAにも、MuBと同様、ssDNAに結合しない状態でdsDNA結合するというモードが存在するということである。IS256*Bsu1*によるmini-ISの転移にRecAまたはMuBが必要である事を踏まえると、MuBとRecAに共通する能力はMutator like トランスポゼース-DNA複合体をゲノム上の別の領域へ誘導する働きであると考えられる。一方、MuBにはRecAのようなDNA修復能は見られない（図12）。*muB*と*recA*は遺伝的に共通の祖先配列を持つ可能性を考慮すると（図8）、相同組換え遺伝子の祖先配列は元来、IS256*Bsu1*や*muA*のなど、DDE型の転移酵素遺伝子の祖先とともにウイルスにコードされていた遺伝子であると考えられる。*muB*にUVによるDNA損

傷を修復する能力が見られなかったことから（図11），相同組換え遺伝子の祖先配列は細胞に取り込まれた後，変異によりDNA修復能を獲得したはずである．今日では，外来遺伝子の獲得には接合伝達や有性生殖などが知られるが，細胞どうしの接触が必要な方法は，相同組換え遺伝子の獲得方法として効率的であるとは考えにくい．ウイルスのような水平伝播のシステムであれば，複製された相同組換え遺伝子の流布を並列して進めることが可能であり，様々な生物種が相同組換え遺伝子を獲得している要因を説明しうる．では，相同組換え遺伝子の祖先配列が実際にはどのような経路で生物に広まったのだろうか．また，DNA修復能を獲得するに至った原動力は何だろうか．

## 2. *recA*祖先配列が相同組換え能を獲得する過程の構築

続いて，相同組換え遺伝子が水平伝播した経路の全容について考える．UVなどの外的要因によるDNA損傷に対応するには，DNA修復因子の存在が必要不可欠であるが，それは生物が地上に進出する上でも同じことである．ここでは，図8に示した系統樹に基づき，初期生命のDNA修復因子の獲得経路の推定を試みた(図17)．

### 2-1. 真核生物と古細菌の有する相同組換え遺伝子について

まず，真核生物と古細菌の相同組換え遺伝子について詳述する．真核生物の祖先細胞（LECA; last eukaryotic common ancestor）の由来については未だ議論の余地があるが[102]，少なくともこの細胞は古細菌のグループであるTACK（Thaumarchaeota, Aigarchaeota, Crenarchaeota, Korarchaeota）を姉妹に持つことは明らかのようなのである[103]．我々の系統解析結果も，*radA*の祖先に近いクレード群はThaumarchaeota/Crenarchaeota/Candidatus Korarchaeotaを含むTACK super familyで構成されており，*rad51*のクレードと隣り合った位置にある(図8，>99%)．従って，真核生物の持つ*rad51*は，TACK superfamilyの*radA*と近縁である．*rad51*のOTUの関係は，一般的な真核生物の進化系統樹と大きく矛盾した点は見当たらず，真核細胞の祖先がもつ*rad51*は単系統(monophyletic group)であると考えられる(図17)．*radA*は，OTU間における配列のGC含量が線型性を示したことを踏まえると（図9. B），古細菌の間で水平伝播により広まった可能性がある．また，*rad51*や*radA*も*recA*と同様，バクテリオファージの相同組換え遺伝子と共通祖先を有するの

だから(図8), 真核生物や古細菌の祖先細胞が*rad51*や*radA*を水平伝播により獲得したと考える方が自然である。以上から, *rad51*や*radA*の祖先配列は, 真核生物と古細菌の共通祖先によって, 水平伝播により獲得された遺伝子であると考えられる。真核生物は*rad51*の祖先配列を獲得した後, 多様化を続け, 今日に至る。

## 2-2. EGTによる*recA*の真核生物への水平伝播

続いて, 真核生物の中でも, 葉緑体とミトコンドリアが細胞共生している陸上植物に注目し, *recA*の真核生物への水平伝播の様子を辿る。陸上植物は他の真核生物とは異なり, 母細胞に由来する*rad51*だけでなく, ミトコンドリア, 葉緑体に由来する*recA*を保存している(このような細胞内共生に由来する遺伝子の水平伝播を, EGT; Endosymbiotic gene transferと呼ぶ)。そのため, 陸上植物の細胞の成り立つプロセスが, これらのクレードの位置関係から構築可能である。図8を確認すると, 陸上植物のミトコンドリア*recA*, 葉緑体*recA*, *rad51*は, 各々一つのクレードを形成している。然るに, 陸上植物の祖先細胞は単系統であり, 祖先となる真核細胞は一度のミトコンドリアとの共生と, それに次いで一度の葉緑体の取り込みがあったことを支持する。熱ショックタンパク質 (Hsp60) のアミノ酸配列を用いた系統解析結果によれば, ミトコンドリアの祖先が $\alpha$ プロテオバクテリア, 葉緑体の祖先がシアノバクテリアに由来すると考えられている[104]。では, ミトコンドリアを獲得した陸上植物の祖先細胞が, 他の真核細胞生物の祖先でもあるのだろうか。陸上植物以外の真核生物が持つ*recA*のOUTの位置を確認すると, 緑藻植物門 (Chlorophyta) の葉緑体*recA*や変形菌 (Mycetozoa) のミトコンドリア*recA*のクレードは, 高等植物細胞由来のものとは枝の位置が異なっていることが分かった(図8)。緑藻植物門の葉緑体*recA*のクレードに近縁のOTUはバクテロイデス門 (Bacteroidetes) やクロロビウム門 (Chlorobi) などであり, シアノバクテリア門やプロテオバクテリア門ではない(クロロビウム門に属す*Chlorobium tepidum*は, シアノバクテリアではないが, 絶対嫌気性の光合成細菌である)。変形菌のミトコンドリア*recA*のクレードは,  $\gamma$ プロテオバクテリア (Gammaproteobacteria) やフェルミクテス門モリクテス綱 (Mollicutes) のクレードに近縁である。つまり, 真核細胞の祖先 (LECA)は*rad51*の系統関係から単系統なのだが, 取り込まれたミトコンドリアや葉緑体については, 多系統(polyphyletic group)である。このことから, 既に*rad51*を獲得して増殖していた真核細胞が, やはり*recA*の祖先配列を有するミトコンドリアや葉緑体の祖先細胞を個別

に取り込んでいたことが分かる。陸上植物のミトコンドリアや葉緑体に由来する*recA*配列が、他のバクテリアに見られるものよりも祖先配列に近いことから、陸上植物の細胞共生は*recA*配列が多様化する前に生じていたと考えられる。植物細胞とは対照的に、ほとんどの真核細胞にミトコンドリア由来の*recA*が見られないが、これは単純に共生の後に不要なために欠損したと考えるのが妥当である。陸上植物がミトコンドリア*recA*を保存しているのは、地上で固着するという生活スタイルから、UVによるDNA損傷の修復を、他の生物よりも高いレベルで維持する必要があったためであると考えられる。例えば、ヒメツリガネゴケ (*Physcomitrella patens*) のミトコンドリア*recA*は、実際にミトコンドリアのDNA損傷を修復していることが確認されている[105]。

以上を簡単にまとめる。まず、真核細胞は既に*rad51*の祖先となる配列を持った状態でミトコンドリアあるいは葉緑体の取り込みを行っていた。つまり、細胞共生が起きた時点は、少なくとも*rad51*獲得の後である。陸上植物の祖先によるミトコンドリアと葉緑体の取り込みは一回、しかも*recA*があまり多様化していない頃に生じたと考えられる。陸上植物以外の真核生物に取り込まれたミトコンドリアと葉緑体は陸上植物のものとは異なる可能性があり、このことは真核生物の細胞共生が独立して起きており、陸上植物は細胞共生が起きた子孫の一系統であることを示している。また、陸上植物とは異なる系統の*recA*を持つ真核生物の存在は、真核細胞の多様化と並行してバクテリアに見られる*recA*遺伝子の多様化と拡散が起きていたことを示唆している。葉緑体を持たない真核生物にはミトコンドリア*recA*が保存されていないため、真核生物の共生の過程を相同組換え遺伝子の系統で構築可能なのはここまでである。

### 2-3. バクテリアにおける*recA*の水平伝播の可能性

では、バクテリアにおける*recA*獲得の様子はどのようなものだったのだろうか。現在、地質学的に確認されている最古のシアノバクテリアの痕跡は21億5千万年前のもの、一方で真核生物の痕跡は17億8千万年から16億8千万年前の間である[106]。しかし、27億年前のドリルコアサンプルから、フィルミクテス (Firmicutes)、放線菌 (Actinobacteria)、各種プロテオバクテリア ( $\alpha$ ,  $\beta$ ,  $\gamma$ ) に属する細菌が発見されている[107]。つまり、27億年前の時点で、これらのバクテリアは一つの生物種として存在していたことを意味する。一方、これらのバ

クテリアの持つ*recA*配列は、ミトコンドリアやシアノバクテリアの物より系統的に新しいため(図8), すべてのバクテリアの祖先が*recA*の祖先配列を持っていたと考えると矛盾が生じる。従って, *recA*の祖先配列は直系に遺伝しておらず, ミトコンドリアやシアノバクテリアの*recA*祖先配列が他のバクテリアに水平伝播したと考える方が自然である(図17). *recA*と*radA*については, 水平伝播によって種間を行き来した痕跡がGC含量の変化から伺える (図18. B,C). この時, 水平伝播に寄与したのは, IS256のようなトランスポゾン, あるいはそのような遺伝子を有するファージであったと考える。ここで特筆すべきは, *recA*クレードの起点にウイルスにコードされる*recA*ホモログが存在することである (図18) . *Microcystis aeruginosa* phage Ma-LMM01(NC\_008562.1)は*Microcystis aeruginosa* (シアノバクテリアの一種) に感染するdsDNAウイルスである[108]. このことは, 相同組換え酵素遺伝子の祖先配列が*recA*に変異する過程にウイルスが介在していることを示す証拠であると考えられる。

### 3. 相同組換え遺伝子の祖先配列が生物に獲得されるプロセス, 及びその生物学的要因.

本節の最後に、紫外線及び宇宙線の遮蔽に先立ち、光合成能が獲得されていることとDNA損傷の抱える矛盾の解消を試み、次いで相同組換え因子がDNA修復能を獲得した原動力とその仕組みについて考える。DNA二本鎖切断の要因となる宇宙線は地球を取り巻く磁場（バンアレン帯）により遮蔽されるが、その形成は今から約28-27億年前であり、紫外線、特にDNAに作用するUV-Cを遮蔽するオゾン層の形成は今から約7-5億年前である[109]。しかし、ストロマトライトは27億年前の地層で産出されている[109]。従って、生物が光合成能を獲得した時期は、紫外線はおろか、宇宙線が遮蔽されるよりも前だということになる。それならば、宇宙線や紫外線に曝されるという環境の下、相同組換え遺伝子を獲得したプロセスはどのように説明できるだろうか。近年、大腸菌RecAの相同組換え能を向上させるという実験が行われた[110]。この実験は、F-の大腸菌に多様な変異を含む*recA*のプラスミドライブラリを形質転換したのち、異なるアミノ酸要求性をもつHfr菌と接合伝達させ、選択培地上に生えた菌からプラスミドを取得するという行程を6度繰り返すというもので、実際に野生型よりも高い相同組換え能を有する大腸菌*recA*の取得に成功している[110]。恐らく、生物が相同組換え因子を獲得した過程も、この実験のような特定の選択条件下における形質転換の繰り返しの中で生じたものであると考える（図19）。つまり、宇宙線や紫外線が選択条件であり、形質転換の方法がバクテリオファージの感染である。しかしながら、以上はDNA修復能のある配列が存在し、強化されるプロセスとも言え、相同組換え遺伝子の祖先配列が”DNA修復能”そのものをどのように獲得したかは説明できない。

MuBにDNA修復能がないこと（図11）、系統解析において*muB*は祖先配列に近いこと（図8）を加味すると、相同組換え遺伝子の祖先配列は本来は相同組換え能を持っていなかったと考えられる。ならば、相同組換え遺伝子の祖先配列が、ウイルスの因子としてではなく、生物自身に利用され始める要因とは一体何だったのだろうか。分子系統解析結果を見ると、*recA*のクレードの起点に、葉緑体やシアノバクテリアよりも先にMollicutesとミトコンドリア（植物）のクレードが存在することが分かる（図8）。これは、Mollicutesやミトコンドリアの祖先細胞の方が、葉緑体の祖先細胞やシアノバクテリアよりも先に*recA*を獲得していることを示唆する。ミトコンドリアはATPを生産するオルガネラであるが、その性質上DNAが酸化損傷を受けやすい。実際、ミトコンドリアにおける突然変異率は核内の染色体よりも高いことが知られている[111]。ここで、相同組換え遺伝子の祖先配列が、元々はMuBの

ようなdsDNA結合活性を有するATPaseをコードしていたする。ATPの過剰な状態では細胞内が酸化的になり、DNA二本鎖切断が生じるが、ATPを結合しdsDNAに結合する因子が細胞内に存在すれば、染色体はこの因子によって二重に保護されることになる。つまり、細胞内に起因するDNA障害の対策として利用されたのが、相同組換え遺伝子の祖先配列が生物にもたらされた最初のきっかけであると考えられるのである。MizunoらはMuBと5つのATPase活性を有するDNA結合型タンパク質（NtrC1, PsPF, ZraR, RFC, DnaA）を比較、相同性を検討しているが[72]、これらの因子をコードする遺伝子を、相同組換え遺伝子と共に分子系統解析を行った場合、*muB*と相同組換え遺伝子の系統関係の間に種々のATPaseを含むDNA結合因子が見出される可能性がある。また、RecAとMuB, UvsXとMuBのアライメント解析結果から、コアドメインに内在するDNA結合領域（loop-L1/L2）と相同なMuBのDNA結合領域は互いに異なることが分かった（図12,13）。また、RecAのC-terminalドメインはDNA結合に関わるが[112,113]、これはRad51には保存されておらず、Rad51のN-terminalドメインもDNA結合に関わるが、この領域はRecAには見られない[114]。これらが示すのは、各相同組換え遺伝子における相同組換え能の獲得は、独立して生じたということである。言い換えれば、これらは収斂進化の結果ということになる。

上述した大腸菌RecAの実験になぞらえるならば [110]、相同組換え酵素遺伝子の変異の様子は次のようになると考えられる。まず、細胞内に過剰に生じたATPを消費し、かつDNAに結合して補強する因子として、バクテリオファージのDNA結合因子が利用されるようになった。その過程で、DNA修復能を獲得した因子が現れ、生物間で利用されるようになった。やがて、生物が光合成能を獲得するに至り、紫外線が届く程度の水深において、相同組換え因子の祖先配列をコードしたウイルスが細胞への感染と溶菌を繰り返す中で、DNA修復能は強化された。この配列を溶原化した細胞は数が増えるので、結果として水平伝播する可能性も高まる。一方で、感染したこの祖先配列と連鎖して遺伝していたファージの遺伝子は、徐々に欠損して機能を失ったと考えられる。何故ならば、ファージが誘導されて溶菌した場合、宿主が死んでしまうため、溶菌しにくいものが自動的に生き残ると考えられるためである。実験結果を踏まえると、この祖先に近い配列が、まるで生きた化石のように*muB*配列に保存されているかも知れない。シロイヌナズナ (*Arabidopsis thaliana*) のミトコンドリア*recA*が、大腸菌の*recA*欠損株におけるUV感受性を完全ではないが相補可能であるという事実は

[115], 生物の生存における遺伝子や塩基配列の重要性は, その由来でなく, そこに内包される機能性にあることを意味する.

以上より, 相同組換え酵素遺伝子の祖先配列における進化プロセスをまとめる.

1. 生物にdsDNA結合型ATPaseがまずもたらされた. その出所はバクテリオファージの複製に関与する因子である.
2. この因子は初めに, 細胞内におけるATPのインフレーションに伴うDNA損傷への対策として利用された.
3. やがて, 光合成能を獲得した生物は, エネルギー獲得と引き換えに宇宙線, 紫外線によるDNA損傷に対応する必要が生じ, 相同組換え酵素遺伝子はDNA修復システムの一部として利用されるようになった.
4. この過程で, 相同組換え遺伝子は宇宙線及び紫外線の下, ファインチューニングされた.
5. dsDNA結合型ATPaseが相同組換え能を獲得したプロセスは, 各相同組換え遺伝子の間で独立であり, 収斂進化の結果である.

筆者が, 相同組換え遺伝子の祖先配列が感染あるいは水平伝播によりもたらされたと考えている要点は, 1.に集約されている.



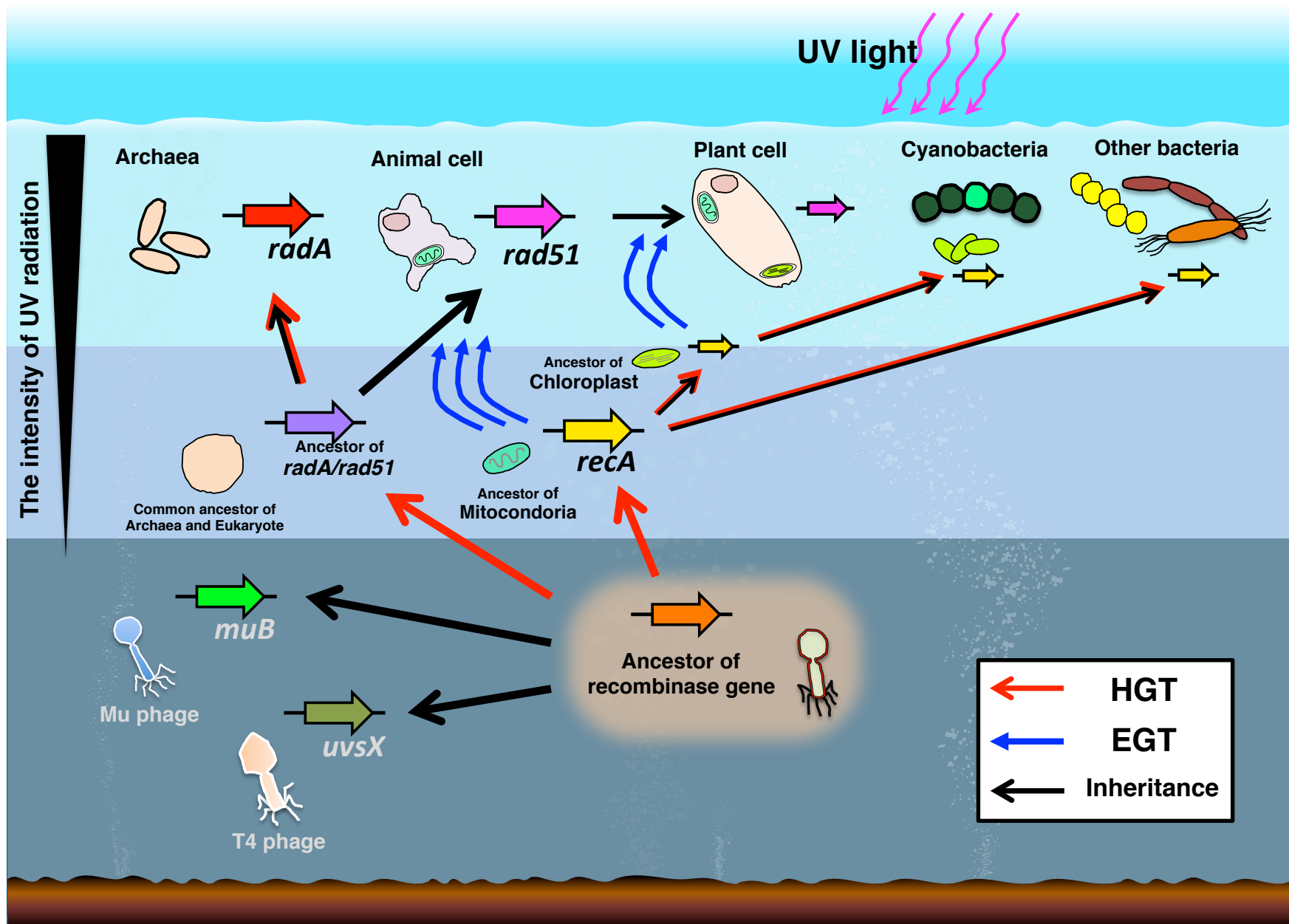


図17. 系統解析結果に基づく *recA* 祖先配列が相同組換え能を獲得する過程の再構築

系統解析結果に基づく、推定可能な相同組換え遺伝子の伝播プロセスを示す。赤矢印は水平伝播（HGT）と考えられる過程，青矢印は細胞共生による遺伝情報の持ち込み（EGT）と考えられる過程，黒矢印は遺伝（Inheritance）を示す。赤黒両矢印が重なっているものについては、水平伝播と遺伝の両方が生じたと考えられる場合を意味する。図の縦軸は、紫外線（UV）の照射強度に対応し、上に行くほど紫外線に暴露する割合が増えることを意味する。黒枠の色付き矢印（i.e. 黄色）は遺伝子を意味している。遺伝子の側のアニメーションは、これらの遺伝子を保存する細胞やバクテリオファージを意味する。

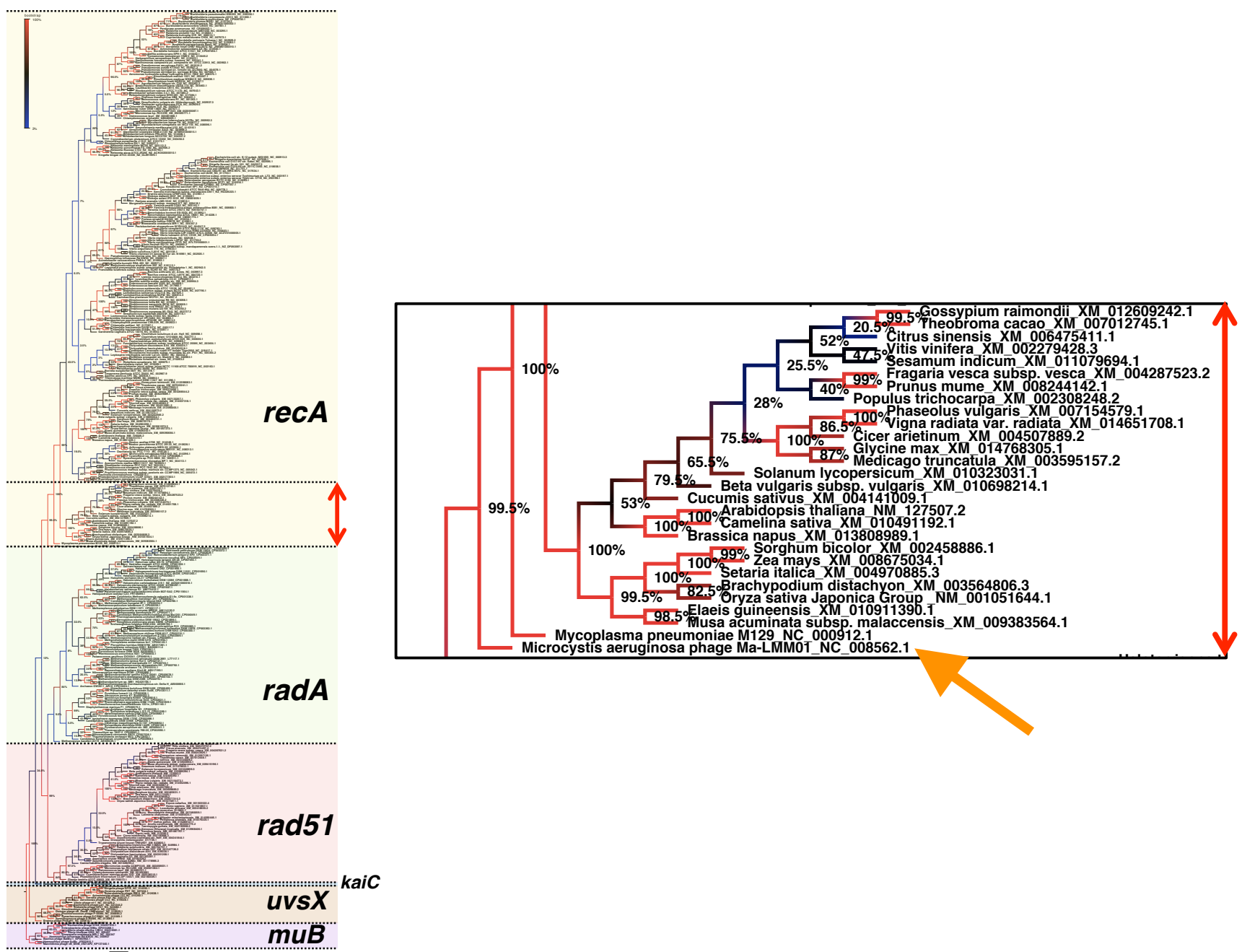


図18. *recA*クレードの起点にウイルスにコードされる*recA*ホモログが存在する

*Microcystis aeruginosa* (シアノバクテリアの一種) に感染するdsDNAウイルス、*Microcystis aeruginosa* phage Ma-LMM01(NC\_008562.1)は相同組換え酵素遺伝子の系統樹における、*recA*クレードの起点に存在する (オレンジ色の矢印) .

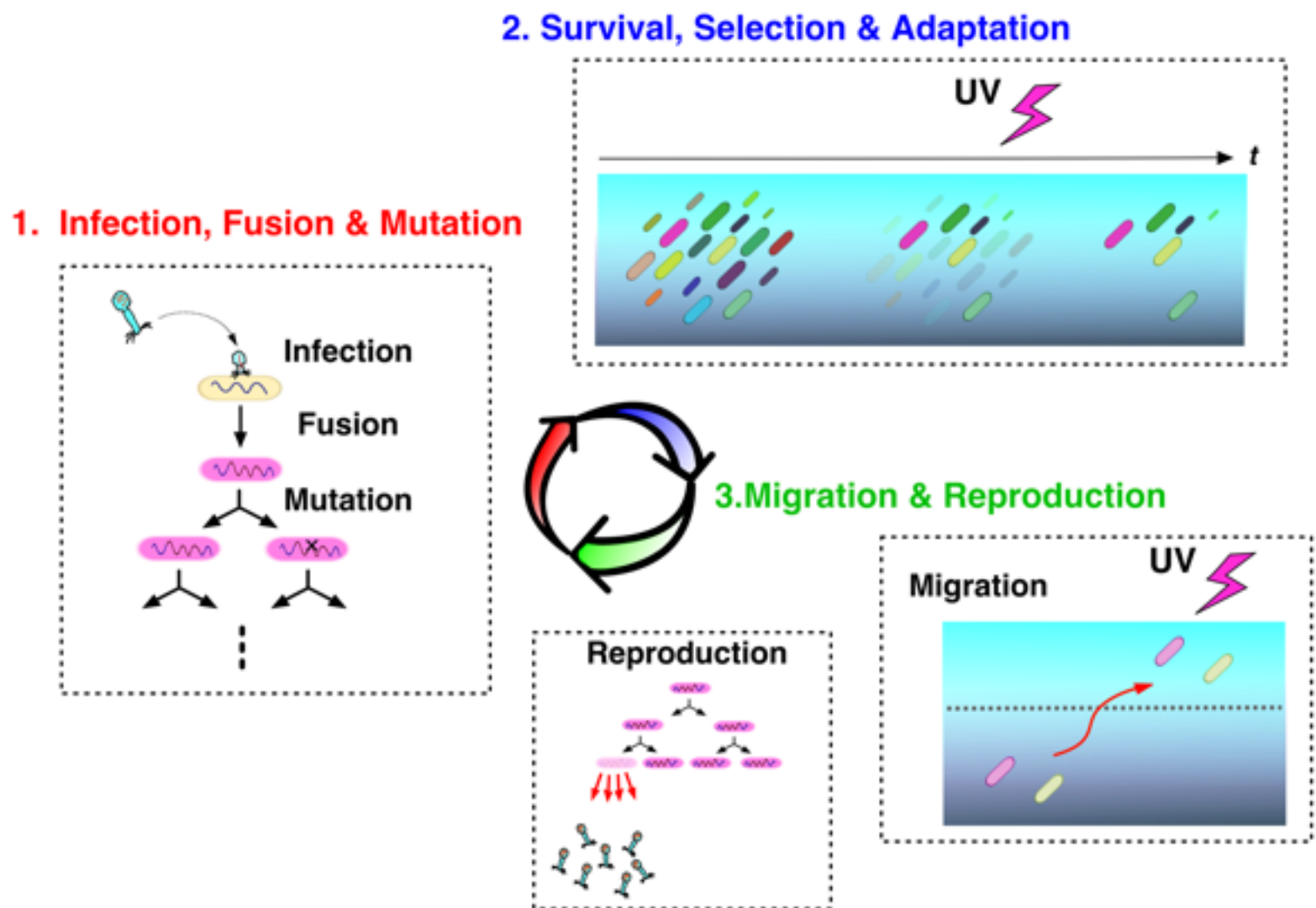


図19. ウイルス感染を介した相同組換え能の進化プロセス.

相同組換え酵素遺伝子の相同組換え能進化プロセスを示す。バクテリオファージ上の相同組換え酵素遺伝子の祖先配列は、感染（Infection）、ウイルスゲノムと宿主細胞のゲノムの融合（Fusion）を介して細胞に取り込まれた後、突然変異（Mutation）により多様化した（1. Infection, Fusion & Mutation）。続いて、紫外線（UV）による選択を受け、生存可能な細胞が生き残った（2. Survival, Selection & Adaptation）。生き残った細胞は、より紫外線の強い環境へと移動し、細胞分裂を行う過程で溶原化していたウイルスが起動、増殖した（3. Migration & Reproduction）。このウイルスゲノム上に存在する相同組換え酵素遺伝子の祖先配列は、初期のものに比べて紫外線耐性付与能の高い因子をコードしている。

## 結論

以上の研究から, *recA*などの相同組換え因子がウイルスにコードされたATPase活性を有するDNA結合タンパク質遺伝子に由来し, 相同組換え酵素として進化したという仮説を得た. Lynn MargulisとDorion Saganにより著された『*Origin of sex: Three Billion Years of Genetic Recombination*』では, 生命の性の獲得, 進化について詳細に議論されているが [116], 取り分け興味深いのは, DNA修復遺伝子が紫外線耐性を与える遺伝子として発達した後, 性をもたらす遺伝子として進化したと論じている点, そしてウイルスの感染を含む様々なレベルの「融合」が多様性を生み出す要因であると考えている点である. 相同組換え酵素遺伝子が後にウイルスによりもたらされたのか, 或いはウイルス感染を介した細胞とウイルスの共進化の過程で相同組換え酵素遺伝子が生じたと考えるのかという点が, 筆者の考察と異なる. また, Margulisらが提示したウイルス感染や細胞共生を含む「融合 (Fusion)」の概念は, 「変異 (Mutation)」と並び立つ, 進化学上重要な表現型変化の様式であると筆者も考える [116]. 相同組換え酵素遺伝子 (*recA*) とバクテリオファージにコードされるDNA結合タンパク質遺伝子 (*muB*) の遺伝的な関連性を示唆した筆者の研究は, 進化の様式における融合の重要性をより強調するものである. ウイルス感染がもたらしたと見られる進化の例は他にも存在し, 例えば, 哺乳類の胎盤獲得に, その種ごとに異なるウイルス感染が関与しており, しかもウイルスの持ち込んだ機能が胎盤形成過程で積極的に利用されている [117]. このような例を鑑みても, 表現型の変化の様式の一つを「融合」とすることは何ら不自然ではない. 真核生物におけるdsDNAウイルスの進化にも見られるように [118], 特定の選択条件 (または環境条件) の下での変異と, 由来のことなるゲノムの融合は, 進化を説明する基本原理の一つであると考えられる. 本研究結果により, 相同組換え因子の祖先がMuファージに近縁のバクテリオファージに由来することが示唆され, 相同組換え遺伝子の進化の様子に具体性を与えることが出来た. これまで述べてきた実験, 解析結果群は, 相同組換え酵素遺伝子が細胞への感染を介して進化したことを示唆する結果であると筆者は結論し, 本論文の締め括りとする.

# 謝辞

東京農業大学, 千葉櫻拓先生に感謝申し上げます。先生には大小様々な負担をお掛けしたと存じます。先生の助け無しに、私が学位論文を仕上げることは有り得ませんでした。本当にありがとうございました。理化学研究所, 柴田武彦先生には、多くの御意見, 御助言賜りましたこと、本当に励みになりました。心より感謝申し上げます。武蔵野大学, 門多真理子先生には、過去のトランスポゾン研究について教えて頂きました。本当にありがとうございました。立教大学, 関根靖彦先生は*muB*と*recA*の実験についてご意見を賜りました。深く感謝致します。立教大学, 河村富士夫先生には、枯草菌を用いた実験生物学の基礎を教えて頂きました。本当にありがとうございました。深く、感謝申し上げます。株式会社ミツカングループ本社中央研究所の加田茂樹さんは、ISと納豆菌の研究について御教授いただきました。ありがとうございました。不均衡進化懇談会にて議論頂きました、株式会社ちとせ研究所の古澤満先生、日立製作所の合田徳夫先生、いわて東北メディカル・メガバンク機構の志波優先生、株式会社ちとせ研究所の笠原堅さん、UTRAの野田正彦さん、ノンフィクションライターの最相葉月さんに御礼申し上げます。古澤先生には近藤宗平博士の書籍をご紹介頂き、本研究を進化学として立ち上げるための礎となりました。深く感謝致します。東京農業大学の上原万里子先生、新藤里絵先生、金子忠一先生、矢嶋俊介先生に御礼申し上げます。私の直面した困難に対する先生方のお力添えなくして、学位論文を執筆することは不可能でした。深く感謝申し上げます。東京農業大学, 生物資源ゲノム解析センターの兼崎友先生には、折に触れて激励の言葉をかけて頂きました。本当にありがとうございました。東京農業大学, 微生物分子遺伝学研究室（現、細胞ゲノム生物学研究室）でお世話になりました、渡辺智先生、大林龍胆さん、児島智子さん、研究室のトランスポゾン解析チーム（ISチーム）でお世話になりました、茂木俊丞さん、吉田智幸さん、原田翔太さん、枯草菌チームの皆様（細村匡太郎さん、多喜乃雄太さん、安藤愛美さん、山下園加さん、北村夏美さん、大坂夏木さん、高松美沙樹さん、高部将宏さん、渡辺正樹さん、徳山麻里さん、美田智也さん）には、私が博士課程入学後、何度も研究の話にお付き合い頂きました。本当にありがとうございました。最後に、様々な形で私に接し、気付きの機会を与え、支え続けてくれた両親、姉、祖父母、叔父母、その他多くの親戚の皆様へ深く感謝申し上げます。ありがとうございました。

平成29年2月24日

明石 基洋

# 参考文献

1. van Zuilen, M. A., Lepland, A., and Arrhenius, G., Reassessing the evidence for the earliest traces of life, *Nature*. 2002;418, 627–630.
2. 近藤宗平, 分子放射線生物学—放射線は生命にどうはたらくか, 東京大学出版会. 1972
3. Sagan C. On the origin and planetary distribution of life. *Radiat Res*. 1961;15(2):174-192.
4. Drees, J. C., Lusetti, S. L., Chitteni-Pattu, S., Inman, R. B., and Cox, M. M. A RecA filament capping mechanism for RecX protein, *Mol. Cell*. 2004;15, 789–798.
5. Shibata, T., DasGupta, C., Cunningham, R. P., and Radding, C. M. Purified *Escherichia coli* recA protein catalyzes homologous pairing of superhelical DNA and single-stranded fragments, *Proc. Natl. Acad. Sci. U. S. A.* 1979;76, 1638–1642.
6. Little, J. W., Edmiston, S. H., Pacelli, L. Z., and Mount, D. W. Cleavage of the *Escherichia coli* lexA protein by the recA protease, *Proc. Natl. Acad. Sci. U. S. A.* 1980;77, 3225–3229.
7. Wojciechowski, M. F., Peterson, K. R., and Love, P. E. Regulation of the SOS response in *Bacillus subtilis*: evidence for a LexA repressor homolog, *J. Bacteriol.* 1991;173, 6489–6498.

8. van der Veen, S., van Schalkwijk, S., Molenaar, D., de Vos, W. M., Abee, T., and Wells-Bennik, M. H. J. The SOS response of *Listeria monocytogenes* is involved in stress resistance and mutagenesis, *Microbiology*. 2010;156, 374–384.
9. Gruenig, M. C., Stohl, E. A., Chitteni-Pattu, S., Seifert, H. S., and Cox, M. M. Less is more: *Neisseria gonorrhoeae* RecX protein stimulates recombination by inhibiting RecA, *J. Biol. Chem.* 2010;285, 37188–37197.
10. Sakai, A. and Cox, M. M. RecFOR and RecOR as distinct RecA loading pathways, *J. Biol. Chem.* 2009;284, 3264–3272.
11. Kidane, D., Carrasco, B., Manfredi, C., et al. Evidence for different pathways during horizontal gene transfer in competent *Bacillus subtilis* cells, *PLoS Genet.* 2009;5, e1000630.
12. Clark, A. J. and Margulies, A. D. Isolation and characterization of recombination-deficient mutants of *Escherichia coli* K12, *Proc. Natl. Acad. Sci. U. S. A.* 1965; 53, 451–459.
13. Lloyd, R. G. and Buckman, C. Conjugational recombination in *Escherichia coli*: genetic analysis of recombinant formation in Hfr × F- crosses, *Genetics*. 1995;139, 1123–1148.
14. Chintapalli, S. V., Bhardwaj, G., Babu, J., et al. Reevaluation of the evolutionary events within recA/RAD51 phylogeny, *BMC Genomics*. 2013;14, 240.
15. Wu, D., Wu, M., Halpern, A., et al. Stalking the fourth domain in metagenomic data: Searching for, discovering, and interpreting novel, deep branches in marker gene phylogenetic trees, *PLoS One*. 2011;6, e18011.
16. Taylor, A. L. Bacteriophage-induced mutation in *Escherichia coli*, *Proc. Natl. Acad. Sci. U. S. A.* 1963;50, 1043–1051.

17. Morgan, G. J., Hatfull, G. F., Casjens, S. and Hendrix, R. W. Bacteriophage Mu genome sequence: analysis and comparison with Mu-like prophages in *Haemophilus*, *Neisseria* and *Deinococcus*, *J. Mol. Biol.* 2002;317, 337–359.
18. Harshey, R. M. The Mu story: how a maverick phage moved the field forward, *Mob. DNA.* 2012; 3, 21.
19. Greene, E. C. and Mizuuchi, K. Target immunity during Mu DNA transposition: transpososome assembly and DNA looping enhance MuA-mediated disassembly of the MuB target complex, *Mol. Cell.* 2002;10, 1367–1378.
20. Mahillon, J. and Chandler, M. Insertion sequences, *Microbiol. Mol. Biol. Rev.* 1998;62, 725–774.
21. Kimura, K. and Itoh, Y. Determination and characterization of IS4Bsu1-insertion loci and identification of a new insertion sequence element of the IS256 family in a natto starter, *Biosci. Biotechnol. Biochem.* 2007;71, 2458–2464.
22. Guérillot, R., Siguier, P., Goubeyre, E., Chandler, M., and Glaser, P. The diversity of prokaryotic DDE transposases of the mutator superfamily, insertion specificity, and association with conjugation machineries, *Genome Biol. Evol.* 2014;6, 260–272.
23. Yuan, Y.-W. and Wessler, SR. The catalytic domain of all eukaryotic cut-and-paste transposase superfamilies, *Proc. Natl. Acad. Sci. U. S. A.* 2011;108, 7884–7889.
24. Eisen, J. A., Benito, M. I., and Walbot, V. Sequence similarity of putative transposases links the maize Mutator autonomous element and a group of bacterial insertion sequences, *Nucleic Acids Res.* 1994;22, 2634–2636.



25. Montañó, S. P., Pigli, Y. Z., and Rice, P. A. The  $\mu$  transpososome structure sheds light on DDE recombinase evolution, *Nature*. 2012;491, 413–417.
26. 大河内 悠貴. 東京農業大学大学院修士論文. 2009.
27. 茂木 俊丞. 東京農業大学大学院修士論文. 2010.
28. Motohiro Akashi, Shota Harada, Syunsuke Moki, Yuki Okouji, Kiwamu Takahashi, Shigeki Kada, Keigo Yamagami, Yasuhiko Sekine, Satoru Watanabe, Taku Chibazakura, and Hirofumi Yoshikawa. Transposition of insertion sequence IS256Bsu1 in *Bacillus subtilis* 168 is strictly dependent on *recA*, *Gene and Genetic Systems*. 2017; in printing.
29. Anagnostopoulos, C. and Spizizen, J. Requirements for transformation in *Bacillus subtilis*, *J. Bacteriol.* 1961;81, 741–746.
30. Leighton, T. J. and Doi, R. H. The stability of messenger ribonucleic acid during sporulation in *Bacillus subtilis*, *J. Biol. Chem.* 1971;246, 3189–3195.
31. Grant, S. G. N., Jessee, J., Bloom, F. R., and Hanahan, D. Differential plasmid rescue from transgenic mouse DNAs into *Escherichia coli* methylation-restriction mutants, *Proc. Natl. Acad. Sci. U. S. A.* 1990;87, 4645-4649.
32. Nishito Y, Osana Y, Hachiya T, et al. Whole genome assembly of a natto production strain *Bacillus subtilis* natto from very short read data. *BMC Genomics*. 2010;11:243.
33. Maamar, H. and Dubnau, D. Bistability in the *Bacillus subtilis* K-state (competence) system requires a positive feedback loop, *Mol. Microbiol.* 2005;56, 615–624.

34. Britton R a, Eichenberger P, Gonzalez-Pastor JE, et al. Genome-wide analysis of the stationary-phase sigma factor (sigma-H) regulon of *Bacillus subtilis*. *J Bacteriol.* 2002;184(17):4881-4890.
35. Siervo, N., Makita, Y., de Hoon, M., and Nakai, K. DBTBS: a database of transcriptional regulation in *Bacillus subtilis* containing upstream intergenic conservation information, *Nucleic Acids Res.* 2008;36 (Database issue), D93–D96.
36. Higuchi, R. Using PCR to engineer DNA, In: Erlich, H. A. (ed)PCR Technology – Principles and Applications for DNA Amplification, *Stockton Press, New York.* 1989; pp. 61–70.
37. Horinouchi S, Weisblum B. Nucleotide sequence and functional map of pE194, a plasmid that specifies inducible resistance to macrolide, lincosamide, and streptogramin type B antibiotics. *J Bacteriol.* 1982;150(2):804-814.
38. Itaya M, Yamaguchi I, Kobayashi K, Endo T, Tanaka T. The blasticidin S resistance gene (bsr) selectable in a single copy state in the *Bacillus subtilis* chromosome. *J Biochem.* 1990;107(6):799-801.
39. Neubert, K. and Brunner, E. A studentized permutation test for the non-parametric Behrens – Fisher problem, *Comput. Stat. Data Anal.* 2007;51, 5192–5204.
40. Jang, S., Sandler, S. J., and Harshey, R. M. Mu insertions are repaired by the double-strand break repair pathway of *Escherichia coli*, *PLoS Genet.* 2012;8, e1002642.
41. Loessner, I., Dietrich, K., Dittrich, D., Hacker, J., and Ziebuhr, W. Transposase-dependent formation of circular IS256 derivatives in *Staphylococcus epidermidis* and *Staphylococcus aureus*, *J. Bacteriol.* 2002;184, 4709–4714.

42. Claverys J-P, Prudhomme M, Martin B. Induction of competence regulons as a general response to stress in gram-positive bacteria. *Annu Rev Microbiol.* 2006;60, 451-475.
43. Hamoen, L. W., Haijema, B., Bijlsma, J. J., Venema, G., and Lovett, C. M. The *Bacillus subtilis* competence transcription factor, ComK, overrides LexA-imposed transcriptional inhibition without physically displacing LexA, *J. Biol. Chem.* 2001;276, 42901–42907.
44. Kurumizaka, H., Ikawa, S., Sarai, A., and Shibata, T. The mutant RecA proteins, RecAR243Q and RecAK245N, exhibit defective DNA binding in homologous pairing, *Arch. Biochem. Biophys.* 1999;365, 83–91.
45. Rice, P., Longden, I., and Bleasby, A. EMBOSS: the European Molecular Biology Open Software Suite, *Trends Genet.* 2000;16, 276–277.
46. Roca, A. I. and Cox, M. M. RecA protein: structure, function, and role in recombinational DNA repair, *Prog. Nucleic Acid Res. Mol. Biol.* 1997;56, 129–223.
47. Britt, R. L., Chitteni-Pattu, S., Page, A. N., and Cox, M. M. RecA K72R filament formation defects reveal an oligomeric RecA species involved in filament extension, *J. Biol. Chem.* 2011;286, 7830–7840.
48. Natri, H. G. and Knight, K. L. Identification of residues in the L1 region of the RecA protein which are important to recombination or coprotease activities, *J. Biol. Chem.* 1994;269, 26311–26322.
49. Natri, H. G., Guzzo, A., Lange, C. S., Walker, G. C., and Knight, K. L. Mutational analysis of the RecA protein L1 region identifies this area as a probable part of the co-protease substrate binding site, *Mol. Microbiol.* 1997;25, 967–978.

50. Shinohara, T., Ikawa, S., Iwasaki, W., et al. Loop L1 governs the DNA-binding specificity and order for RecA-catalyzed reactions in homologous recombination and DNA repair, *Nucleic Acids Res.* 2015;43, 973–986.
51. Hortnagel, K., Voloshin, O. N., Kinal, H. H., Ma, N., Schaffer-Judge, C., and Camerini-Otero, R. D. Saturation mutagenesis of the E. coli RecA loop L2 homologous DNA pairing region reveals residues essential for recombination and recombinational repair, *J. Mol. Biol.* 1999;286, 1097–1106.
52. De Zutter, J. K., Forget, A. L., Logan, K. M., and Knight, K. L. Phe217 regulates the transfer of allosteric information across the subunit interface of the RecA protein filament, *Structure.* 2001;9, 47–55.
53. Skiba, M. C. and Knight, K. L. Functionally important residues at a subunit interface site in the RecA protein from Escherichia coli, *J. Biol. Chem.* 1994;269, 3823–3828.
54. Lee, C.-D. and Wang, T.-F. The N-terminal domain of Escherichia coli RecA have multiple functions in promoting homologous recombination, *J. Biomed. Sci.* 2009;16, 37.
55. Manfredi C, Carrasco B, Ayora S, Alonso JC. Bacillus subtilis RecO nucleates RecA onto SsbA-coated single-stranded DNA. *J Biol Chem.* 2008;283(36):24837-24847.
56. Adzuma, K. and Mizuuchi, K. Target immunity of Mu transposition reflects a differential distribution of Mu B protein, *Cell.* 1988;53, 257–266.
57. Krukltis, R., Welty, D. J., and Nakai, H. ClpX protein of Escherichia coli activates bacteriophage Mu transposase in the strand transfer complex for initiation of Mu DNA synthesis, *EMBO J.* 1995;9, 935–944.

58. Levchenko, I., Yamauchi, M., and Baker, T. A. ClpX and MuB interact with overlapping regions of Mu transposase: implications for control of the transposition pathway, *Genes Dev.* 1997;11, 1561–1572.
59. Greene, E. C. and Mizuuchi, K. Visualizing the assembly and disassembly mechanisms of the MuB transposition targeting complex, *J. Biol. Chem.* 2004;279, 16736–16743.
60. Muller, B., Koller, T., and Stasiak, A. Characterisation of the DNA binding activity of stable RecA-DNA complexes: interactions between the two DNA binding sites within RecA helical filaments, *J. Mol. Biol.* 1990;212, 97–112.
61. Conover, A. J., Danilowicz, C., Gunaratne, R., Coljee, V. W., Kleckner, N., and Prentiss, M. Changes in the tension in dsDNA alter the conformation of RecA bound to dsDNA-RecA filaments, *Nucleic Acids Res.* 2011;39, 8833–8843.
62. Bianco, P. R., Tracy, R. B., and Kowalczykowski, S. C. DNA strand exchange proteins: a biochemical and physical comparison, *Front. Biosci.* 1998;3, D570–D603.
63. Haldenby, S., White, M. F., and Allers, T. RecA family proteins in archaea: RadA and its cousins, *Biochem. Soc. Trans.* 2009;37, 102–107.
64. Hasegawa, Y., Wakabayashi, M., Nakamura, S., Kodaira, K., Shinohara, H., and Yasukawa, H. A homolog of Escherichia coli RecA in mitochondria of the cellular slime mold Dictyostelium discoideum, *DNA Repair.* 2004;3, 515–525.
65. Katoh, K. and Standley, D. M. MAFFT multiple sequence alignment software version 7: improvements in performance and usability, *Mol. Biol. Evol.* 2013;30, 772–780.
66. Chang, J., Di Tommaso, P., and Notredame, C. TCS: a new multiple sequence alignment reliability measure to estimate alignment accuracy and improve phylogenetic tree reconstruction, *Mol. Biol. Evol.* 2014;31, 1625–1637.

67. Nei, M. and Kumar, S. *Molecular Evolution and Phylogenetics*, Oxford University Press, New York. 2000.
68. Felsenstein, J. Confidence limits on phylogenies: an approach using the bootstrap, *Evolution*, 1985;39, 783-791.
69. Kumar, S., Stecher, G., and Tamura, K. MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets, *Mol. Biol. Evol.* 2016;33, 1870-1874.
70. Nishida, H. Evolution of genome base composition and genome size in bacteria, *Front. Microbiol.* 2012;3, 1–3.
71. Mustard, J. A. and Little, J. W. Analysis of Escherichia coli RecA interactions with LexA, lambda CI, and UmuD by site-directed mutagenesis of recA, *J. Bacteriol.* 2000;182, 1659–1670.
72. Mizuno N, Dramićanin M, Mizuuchi M, et al. MuB is an AAA+ ATPase that forms helical filaments to control target selection for DNA transposition. *Proc Natl Acad Sci U S A.* 2013;110(27):E2441-50.
73. Yonesaki, T., Minagawa, T. T4 phage gene uvsX product catalyzes homologous DNA pairing. *EMBO J.* 1985 Dec 1;4(12):3321-7.
74. Salinas, F., Kodadek, T. Phage T4 homologous strand exchange: a DNA helicase, not the strand transferase, drives polar branch migration. *Cell.* 1995 Jul 14;82(1):111-9.
75. Minagawa, T., Fujisawa, H., Yonesaki, T., Ryo, Y. Function of cloned T4 recombination genes, uvsX and uvsY, in cells of Escherichia coli. *Mol Gen Genet.* 1988 Feb;211(2):350-6.

76. Lusetti SL, Cox MM. The bacterial RecA protein and the recombinational DNA repair of stalled replication forks. *Annu Rev Biochem.* 2002;71:71-100.
77. Gajewski S, Webb MR, Galkin V, Egelman EH, Kreuzer KN, White SW. Crystal structure of the phage T4 recombinase UvsX and its functional interaction with the T4 sf2 helicase uvsw. *J Mol Biol.* 2011;405(1):65-76.
78. Luria SE, Delbrück M. Mutations of Bacteria from Virus Sensitivity to Virus Resistance. *Genetics.* 1943;28(6):491-511.
79. Mccarty M. From the Hospital of The Rockefeller Institute for Medical Research. *J Exp Med.* 1944;79(6):137-158.
80. HERSHEY AD, CHASE M. Independent functions of viral protein and nucleic acid in growth of bacteriophage. *J Gen Physiol.* 1952;36(1):39-56.
81. WATSON JD, CRICK FH. Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. *Nature.* 1953;171(4356):737-738.
82. J Leszczynski. Computational Chemistry: Reviews of Current Trends. *World Scientific Pub Co Inc.* 1997
83. ニコラス・H. バートン, ジョナサン・A. アイゼン, デイビッド・B. ゴールドステイン, ニパム・H. パテル, デレク・E.G. ブリッグス 著, 宮田 隆, 星山 大介 翻訳., 2009 「進化—分子・個体・生態系」, メディカルサイエンスインターナショナル
84. Peng CS, Fedeles BI, Singh V, et al. Two-dimensional IR spectroscopy of the anti-HIV agent KP1212 reveals protonated and neutral tautomers that influence pH-dependent mutagenicity. *Proc Natl Acad Sci.* 2015;112(11):201415974.

85. Minnick DT, Liu L, Grindley NDF, Kunkel TA, Joyce CM. Discrimination against purine-pyrimidine mispairs in the polymerase active site of DNA polymerase I: a structural explanation. *Proc Natl Acad Sci U S A*. 2002;99(3):1194-1199.
86. Akashi M, Yoshikawa H. Relevance of GC content to the conservation of DNA polymerase III/mismatch repair system in Gram-positive bacteria. *Front Microbiol*. 2013;4(September):266.
87. Boratyn GM, Schäffer A a, Agarwala R, Altschul SF, Lipman DJ, Madden TL. Domain enhanced lookup time accelerated BLAST. *Biol Direct*. 2012;7:12.
88. 松下貢著, フラクタルの物理 〈1〉 基礎編, 裳華房. 2002.
89. Durbin RM, Altshuler DL, Durbin RM, et al. A map of human genome variation from population-scale sequencing. *Nature*. 2010;467(7319):1061-1073.
90. Nagasaki M, Yasuda J, Katsuoka F, et al. Rare variant discovery by deep whole-genome sequencing of 1,070 Japanese individuals. *Nat Commun*. 2015;6:8018.
91. Hilbert D. Ueber die stetige Abbildung einer Linie auf ein Flaechenstueck. *Math Ann*. 1891;38(3):459-460.
92. Lieberman-Aiden E, van Berkum NL, Williams L, et al. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science*. 2009;326(October):289-293.
93. Mirny L a. The fractal globule as a model of chromatin architecture in the cell. *Chromosom Res*. 2011;19:37-51.



94. Gouda N, Shiwa Y, Akashi M, Yoshikawa H, Kasahara K, Furusawa M. Distribution of human single-nucleotide polymorphisms is approximated by the power law and represents a fractal structure. *Genes to Cells*. 2016;21(5):396-407.
95. Martincorena I, Seshasayee ASN, Luscombe NM. Evidence of non-random mutation rates suggests an evolutionary risk management strategy. *Nature*. 2012;485(7396):95-98.
96. Hurst LD, Merchant AR. High guanine-cytosine content is not an adaptation to high temperature: a comparative analysis amongst prokaryotes. *Proc Biol Sci*. 2001;268(1466):493-497.
97. Kawai Y, Maeda Y. GC-content of tRNA genes classifies archaea into two groups. *J Gen Appl Microbiol*. 2009;55(5):403-408.
98. Epstein B, Jones M, Hamede R, et al. Rapid evolutionary response to a transmissible cancer in Tasmanian devils. *Nat Commun*. 2016;7:12684.
99. 木村 資生, 向井 輝美, 日下部 真一著, 分子進化の中立説, 紀伊國屋書店. 1986.
100. Kimura M. Evolutionary rate at the molecular level. *Nature*. 1968;217(5129):624-626.
101. Akashi H, Osada N, Ohta T. Weak selection and protein evolution. *Genetics*. 2012;192(1):15-31.
102. López-García, P and Moreira, D. Open Questions on the Origin of Eukaryotes, *Trends Ecol. Evol*. 2015;30, 697–708.

103. Koonin, E. V. Origin of eukaryotes from within archaea, archaeal eukaryome and bursts of gene gain: eukaryogenesis just made easier? *Phil. Trans. R. Soc. B.* 2015;370, 20140333.
104. Viale, A. M. and Arakaki, A. K. The chaperone connection to the origins of the eukaryotic organelles, *FEBS Lett.* 1994;341, 146–151.
105. Odahara, M., Inouye, T., Fujita, T., Hasebe, M., and Sekine, Y. Involvement of mitochondrial-targeted RecA in the repair of mitochondrial DNA in the moss, *Physcomitrella patens*, *Genes Genet. Syst.* 2007;82, 43–51.
106. Rasmussen, B., Fletcher, I. R., Brocks, J. J., and Kilburn, M. R. Reassessing the first appearance of eukaryotes and cyanobacteria, *Nature.* 2008;455, 1101–1104.
107. Gérard, E., Moreira, D., Philippot, P., Van Kranendonk, M. J., and López-García, P. Modern subsurface bacteria in pristine 2.7 Ga-old fossil stromatolite drillcore samples from the fortescue group, western Australia, *PLoS One.* 2009;4, e5298.
108. Yoshida T, Nagasaki K, Takashima Y, et al. Ma-LMM01 infecting toxic *Microcystis aeruginosa* illuminates diverse cyanophage genome strategies. *J Bacteriol.* 2008;190(5):1762-1772.
109. 丸山茂徳, 磯崎行雄, 生命と地球の歴史, 岩波書店. 1998.
110. Kim, T., Chitteni-Pattu, S., Cox, B. L., Wood, E. A., Sandler, S. J., and Cox, M. M. Directed evolution of RecA variants with enhanced capacity for conjugational recombination, *PLoS Genet.* 2015;11, e1005278.

111. Linnane AW, Marzuki S, Ozawa T, Tanaka M. Mitochondrial DNA mutations as an important contributor to ageing and degenerative diseases. *Lancet (London, England)*. 1989;1(8639):642-645.
112. Kurumizaka H, Aihara H, Ikawa S, et al. A possible role of the C-terminal domain of the RecA protein. A gateway model for double-stranded DNA binding. *J Biol Chem*. 1996;271(52):33515-33524.
113. Aihara H, Ito Y, Kurumizaka H, Terada T, Yokoyama S, Shibata T. An interaction between a specified surface of the C-terminal domain of RecA protein and double-stranded DNA for homologous pairing. *J Mol Biol*. 1997;274(2):213-221.
114. Aihara H, Ito Y, Kurumizaka H, Yokoyama S, Shibata T. The N-terminal domain of the human Rad51 protein binds DNA: structure and a DNA binding surface as revealed by NMR. *J Mol Biol*. 1999;290(2):495-504.
115. Khazi, F. R., Edmondson, A. C., and Nielsen, B. L. An Arabidopsis homologue of bacterial RecA that complements an E. coli recA deletion is targeted to plant mitochondria, *Mol. Genet. Genomics*. 2003;269, 454–463.
116. Lynn Margulis, Dorion Sagan. Origin of sex: Three Billion Years of Genetic Recombination, Yale University Press. 1982.
117. Imakawa K, Nakagawa S, Miyazawa T. Baton pass hypothesis: Successive incorporation of unconserved endogenous retroviral genes for placentation during mammalian evolution. *Genes to Cells*. 2015;20(10):771-788.
118. Koonin, E. V., Krupovic, M., and Yutin, N. Evolution of double-stranded DNA viruses of eukaryotes: from bacteriophages to transposons to giant viruses, *Ann. N. Y. Acad. Sci*. 2015;1341, 10–24.

# 付録

付録 1: 検出された重複配列周辺の塩基配列

付録 2: 分子系統解析に用いた相同組換え酵素遺伝子リスト

付録 3: 相同組換え酵素遺伝子の系統樹 詳細

付録 4: 分子系統解析と塩基配列のGC含量の関係性解析

付録 5: 株の作製に用いたプライマーのリスト及び作製した株の構造

付録 6: 実験に用いた菌株リスト

付録 7: 転移頻度測定結果表

付録 8: ゲノム構造と高頻度変異領域の関係性を探るためのシミュレーション解析

付録 9: 培地, 試薬, プロトコール

# 検出された重複配列周辺の塩基配列

第2章で示したmini-ISの転移に伴い検出された重複配列周辺の塩基配列を示す.

Start Position	5' -> 3'	End position	Inserted gene
44839	AAGAAATGATTATTTAAGGT <b>TTTGAAGC</b> TGGTTTTGGATTTTCGCTGAT	44886	<i>yabC</i>
1441388	ACGTGCTGGAATCCTTCCCT <b>AATAAACAT</b> GTAAACCGTGATTACTTTG	1441435	<i>queF</i>
2552999	GTGCCTGAGCAGAGGCACTA <b>ACTCCTCT</b> TTTTGTCAATAACCATCGAGA	2553046	<i>tasA</i>
3109562	GTGGAAGAGGCCGACAACGA <b>GCGGAATG</b> CCGAAGTTGTACATTTGTTT	3109609	<i>bceB</i>
3255525	GTAATAAATCAAATCTTT <b>GAGATGAA</b> TCTTTTTTCCACTTCTCAA	3255572	<i>comP</i>
3785744	ACGGACTGTGACAAAGAGCC <b>AAATGCAT</b> TTTGAATCAGCTCTTTTTTC	3785791	<i>atpH</i>

### 付録 1 表. 検出された重複配列の上流/下流20 bpの配列.

mini-ISの転移により生じた重複配列の上流/下流20 bpの配列を示す (5' -> 3') . 赤いフォントで示した配列は検出した重複配列である. “Start position”及び “End position”は、解読された枯草菌ゲノム (GenBank: CP010052.1) においてアノテーションされた座標と対応する. , Inserted geneはmini-ISが挿入された遺伝子を示す.

# 分子系統解析に用いた 相同組換え酵素遺伝子リスト

第3章で相同組換え酵素遺伝子の系統樹の作製に使用した遺伝子のID, そのGC含量, 遺伝子を保持する生物種名とそのID, 及びその生物種のゲノムのGC含量のリストを掲載する.

Gene	Organism	Taxonomy D	Lineage (superkingdom)	Lineage (phylum)	Lineage (class)	Host genome GC content(%)	Gene GC content(%)	Accession ID	Start position (bp)	End position (bp)	Length (bp)
kaiC	Halosimplex carlsbadense 2-9-1	797114	Archaea	Euryarchaeota	Halobacteria	67.7	68.2	NZ_AOIU01000018.1	12005	13468	1464
muB	Haemophilus influenzae Rd KW20	71421	Bacteria	Proteobacteria	Gammaproteobacteria	38.0	47.7	NC_000907	1564041	1564904	864
muB	Shewanella oneldensis MR-1	211586	Bacteria	Proteobacteria	Gammaproteobacteria	45.9	51.2	NC_004347	2791500	2792456	957
muB	Vibrio cholerae 0395	345073	Bacteria	Proteobacteria	Gammaproteobacteria	47.5	54.6	NC_009457	692300	693244	945
muB	Bacteriophage Mu	10677	Viruses	dsDNA viruses, no RNA stage	Caudovirales	52.1	48.3	AF083977.1	3358	4296	939
muB	Enterobacteria phage SfMu	1567022	Viruses	dsDNA viruses, no RNA stage	Caudovirales	51.9	48.0	KP010268.1	3296	4234	939
muB	Escherichia phage D108	665033	Viruses	dsDNA viruses, no RNA stage	Caudovirales	51.8	48.5	GQ357916.1	3306	4244	939
muB	Haemophilus phage SuMu	483266	Viruses	dsDNA viruses, no RNA stage	Caudovirales	41.9	40.9	JF832915.1	4039	4920	882
muB	Mannheimia phage vB_MhM_3927AP2	1572750	Viruses	dsDNA viruses, no RNA stage	Caudovirales	43.1	44.4	KP137440.1	3330	4211	882
muB	Vibrio phage martha 12B12	573175	Viruses	dsDNA viruses, no RNA stage	unclassified dsDNA phages	45.8	48.5	HQ316581.1	28395	29342	948
muB	Bacillus phage BalMu-1	1567487	Viruses	unclassified phages	unclassified phages	42.8	44.7	KP063902.1	5170	6069	900
rad51	Plasmodium falciparum strain 3D7	36329	Eukaryota	Alveolata	Apicomplexa	22.9	34.1	XM_001347726.2	1	1053	1053
rad51	Entamoeba histolytica HM-1:IMSS	294381	Eukaryota	Amoebozoa	Archamoebae	24.3	33.6	XM_648984.1	1	1101	1101
rad51	Acanthamoeba castellanii str. Neff	1257118	Eukaryota	Amoebozoa	Discosea	57.8	60.5	XM_004341844.1	1	1029	1029
rad51	Dictyostelium discoideum AX4	352472	Eukaryota	Amoebozoa	Mycetozoa	22.5	34.5	XM_639330.1	1	1056	1056
rad51	Dictyostelium fasciculatum	261658	Eukaryota	Amoebozoa	Mycetozoa	33.8	41.3	XM_004361208.1	1	1080	1080
rad51	Trypanosoma brucei brucei TREU927	185431	Eukaryota	Euglenozoa	Kinetoplastida	46.8	50.9	XM_823800.1	1	1122	1122
rad51	Giardia lamblia ATCC 50803	184922	Eukaryota	Fornicata	Diplomonadida	49.2	52.6	XM_001709373.1	1	1170	1170
rad51	Aspergillus oryzae RIB40	510516	Eukaryota	Opisthokonta	Fungi	48.3	49.9	XM_001825262	106	1152	1047
rad51	Saccharomyces cerevisiae S288c	559292	Eukaryota	Opisthokonta	Fungi	38.4	46.7	NM_001178986.3	1	1203	1203
rad51	Alligator mississippiensis	8496	Eukaryota	Opisthokonta	Metazoa	44.4	48.2	XM_014595446.1	315	1334	1020
rad51	Anolis carolinensis	28377	Eukaryota	Opisthokonta	Metazoa	40.8	45.5	XM_003224775.2	278	1297	1020
rad51	Caenorhabditis elegans	6239	Eukaryota	Opisthokonta	Metazoa	35.4	44.9	NM_001028295.2	1	1074	1074
rad51	Ciona intestinalis	7719	Eukaryota	Opisthokonta	Metazoa	36.1	43.0	XM_002126898.3	64	1080	1017
rad51	Danio rerio	7955	Eukaryota	Opisthokonta	Metazoa	36.9	51.7	NM_213206.2	30	1052	1023
rad51	Drosophila melanogaster	7227	Eukaryota	Opisthokonta	Metazoa	41.9	57.6	D17726.1	63	1073	1011
rad51	Equus caballus	9796	Eukaryota	Opisthokonta	Metazoa	41.5	47.7	XM_001503522.4	360	1379	1020
rad51	Gallus gallus	9031	Eukaryota	Opisthokonta	Metazoa	42.9	47.7	XM_015286147.1	183	1202	1020
rad51	Homo sapiens	9606	Eukaryota	Opisthokonta	Metazoa	41.3	47.7	XM_011521857.1	50	1069	1020
rad51	Latimeria chalumnae	7897	Eukaryota	Opisthokonta	Metazoa	42.6	51.0	XM_014485434.1	200	1129	930
rad51	Loxodonta africana	9785	Eukaryota	Opisthokonta	Metazoa	40.9	49.6	XM_003418655.2	242	1261	1020
rad51	Monodelphis domestica	13616	Eukaryota	Opisthokonta	Metazoa	38.1	49.7	XM_007480059.1	21	1040	1020
rad51	Mus musculus	10090	Eukaryota	Opisthokonta	Metazoa	42.1	48.7	D13803.1	189	1208	1020
rad51	Pelodiscus sinensis	13735	Eukaryota	Opisthokonta	Metazoa	44.5	46.9	XM_014576320.1	28	1047	1020
rad51	Taeniopygia guttata	59729	Eukaryota	Opisthokonta	Metazoa	41.5	56.3	XM_002195968.2	1	1020	1020
rad51	Xenopus (Silurana) tropicalis	8364	Eukaryota	Opisthokonta	Metazoa	40.5	49.3	XM_012968406.1	113	1123	1011
rad51	Xenopus laevis	8355	Eukaryota	Opisthokonta	Metazoa	39.5	48.9	NM_001087767.1	164	1174	1011
rad51	Trichomonas vaginalis G3	412133	Eukaryota	Parabasalia	Trichomonadida	32.9	42.3	XM_001303201.1	1	990	990
rad51	Bigeloviella natans	227086	Eukaryota	Rhizaria	Cercozoa	44.9	31.0	XM_001712896.1	1	996	996
rad51	Cyanidioschyzon merolae strain 10D	280699	Eukaryota	Rhodophyta	Bangiophyceae	55.0	57.3	XM_005538310.1	1	1044	1044
rad51	Galdieria sulphuraria	130081	Eukaryota	Rhodophyta	Bangiophyceae	37.9	39.4	XM_005704737.1	5	1102	1098
rad51	Phaeodactylum tricornutum CCAP 1055/1	556484	Eukaryota	Stramenopiles	Bacillariophyta	48.8	54.1	XM_002180409.1	33	1085	1053
rad51	Chlamydomonas reinhardtii	3055	Eukaryota	Viridiplantae	Chlorophyta	62.0	63.9	XM_001692681	263	1294	1032
rad51	Micromonas pusilla CCMP1545	564608	Eukaryota	Viridiplantae	Chlorophyta	65.7	65.5	XM_003058831.1	203	1438	1236
rad51	Micromonas sp. RCC299	296587	Eukaryota	Viridiplantae	Chlorophyta	61.5	63.2	XM_002501362.1	1	1035	1035
rad51	Ostreococcus tauri	70448	Eukaryota	Viridiplantae	Chlorophyta	59.2	62.4	XM_003080701.1	1	1263	1263
rad51	Arabidopsis thaliana	3702	Eukaryota	Viridiplantae	Streptophyta	36.7	45.7	NM_122092.2	92	1120	1029
rad51	Beta vulgaris subsp. vulgaris	3555	Eukaryota	Viridiplantae	Streptophyta	36.5	44.2	XM_010689294.1	183	1211	1029
rad51	Brachypodium distachyon	15368	Eukaryota	Viridiplantae	Streptophyta	46.4	50.2	XM_003577315.3	88	1110	1023



rad51	Brassica napus	Eukaryota	Viridiplantae	Streptophyta	37.6	47.1	XM_013810440.1	87	1115	1029
rad51	Camelina sativa	Eukaryota	Viridiplantae	Streptophyta	37.4	45.8	XM_010456092.1	77	1105	1029
rad51	Cicer arietinum	Eukaryota	Viridiplantae	Streptophyta	32.4	43.9	XM_004507992.2	217	1239	1023
rad51	Citrus sinensis	Eukaryota	Viridiplantae	Streptophyta	36.4	44.8	XM_006474497.2	114	1142	1029
rad51	Cucumis sativus	Eukaryota	Viridiplantae	Streptophyta	36.1	45.7	XM_004144009.2	147	1169	1023
rad51	Elaeis guineensis	Eukaryota	Viridiplantae	Streptophyta	41.0	46.2	XM_010938950.1	212	1240	1029
rad51	Fragaria vesca subsp. vesca	Eukaryota	Viridiplantae	Streptophyta	38.9	47.2	XM_004287631.2	291	1304	1014
rad51	Glycine max	Eukaryota	Viridiplantae	Streptophyta	35.0	46.7	XM_003550987.3	172	1203	1032
rad51	Gossypium raimondii	Eukaryota	Viridiplantae	Streptophyta	33.7	46.1	XM_012597136.1	156	1187	1032
rad51	Medicago truncatula	Eukaryota	Viridiplantae	Streptophyta	34.0	44.2	XM_003609889.2	245	1270	1026
rad51	Musa acuminata subsp. malaccensis	Eukaryota	Viridiplantae	Streptophyta	40.7	46.0	XM_009415196.1	74	1102	1029
rad51	Oryza sativa Japonica Group	Eukaryota	Viridiplantae	Streptophyta	43.5	49.2	NM_001074811.1	90	1109	1020
rad51	Phaseolus vulgaris	Eukaryota	Viridiplantae	Streptophyta	36.9	48.2	XM_007154473.1	171	1205	1035
rad51	Populus trichocarpa	Eukaryota	Viridiplantae	Streptophyta	34.1	44.5	XM_006381487.1	85	1131	1047
rad51	Prunus mume	Eukaryota	Viridiplantae	Streptophyta	38.3	46.4	XM_008243926.1	205	1227	1023
rad51	Sesamum indicum	Eukaryota	Viridiplantae	Streptophyta	35.2	45.4	XM_011079852.1	73	1098	1026
rad51	Setaria italica	Eukaryota	Viridiplantae	Streptophyta	46.2	49.9	XM_004979669.2	52	1074	1023
rad51	Solanum lycopersicum	Eukaryota	Viridiplantae	Streptophyta	35.7	43.7	NM_001246859.3	89	1117	1029
rad51	Sorghum bicolor	Eukaryota	Viridiplantae	Streptophyta	44.3	49.6	XM_002463031.1	1	1035	1035
rad51	Theobroma cacao	Eukaryota	Viridiplantae	Streptophyta	35.2	46.1	XM_007012448.1	369	1454	1086
rad51	Vigna radiata var. radiata	Eukaryota	Viridiplantae	Streptophyta	34.1	45.6	XM_014652096.1	135	1169	1035
rad51	Vitis vinifera	Eukaryota	Viridiplantae	Streptophyta	35.0	45.0	XM_002273767.2	139	1152	1014
rad51	Zea mays	Eukaryota	Viridiplantae	Streptophyta	46.8	48.6	NM_001111448.1	53	1075	1023
radA	Nanoarchaeota archaeon 7A	Archaea	DPANN group	Nanoarchaeota	24.2	32.7	CP010514.1	252206	253183	978
radA	Nanoarchaeum equitans Kin4-M	Archaea	DPANN group	Nanoarchaeota	31.6	37.2	AE017199.1	375398	376375	978
radA	Archaeoglobus fulgidus DSM 4304	Archaea	Euryarchaeota	Archaeoglobi	48.6	49.2	AE000782.1	894676	895689	1014
radA	Ferroglobus placidus DSM 10642	Archaea	Euryarchaeota	Archaeoglobi	44.1	47.6	CP001899.1	1237944	1238966	1023
radA	Geoglobus acetivorans strain SBH6	Archaea	Euryarchaeota	Archaeoglobi	46.8	48.7	CP009552.1	1449814	1450854	1041
radA	Aciduliprofundum boonei T469	Archaea	Euryarchaeota	DHVE2 group	39.2	44.0	CP001941.1	426732	427706	975
radA	Halalkalicoccus jeotgall B3	Archaea	Euryarchaeota	Halobacteria	62.6	65.9	CP002062.1	268659	269693	1035
radA	Halanaeroarchaeum sulfurreducens strain M27-SA2	Archaea	Euryarchaeota	Halobacteria	62.8	65.1	CP011564.1	2065188	2066252	1065
radA	Haloarcula marismortui ATCC 43049	Archaea	Euryarchaeota	Halobacteria	61.1	62.8	AY596297.1	2577913	2578968	1056
radA	Halobacterium salinarum R1	Archaea	Euryarchaeota	Halobacteria	65.8	65.2	AM774415.1	1838012	1839043	1032
radA	Haloferox volcanii DS2	Archaea	Euryarchaeota	Halobacteria	65.6	64.4	CP001956.1	101807	102838	1032
radA	Halogeometricum borinquense DSM 11551	Archaea	Euryarchaeota	Halobacteria	59.9	60.9	CP001690.1	99578	100609	1032
radA	Halomicrobium mukohataei DSM 12286	Archaea	Euryarchaeota	Halobacteria	65.5	65.6	CP001688.1	1123769	1124818	1050
radA	Halopiger xanaduensis SH-6	Archaea	Euryarchaeota	Halobacteria	65.2	66.3	CP002839.1	602401	603432	1032
radA	Haloquadratum walsbyi C23	Archaea	Euryarchaeota	Halobacteria	47.9	46.8	FR746099.1	2949472	2950503	1032
radA	Halorhabdus utahensis DSM 12940	Archaea	Euryarchaeota	Halobacteria	62.9	64.3	CP001687.1	1245090	1246136	1047
radA	Halorubrum lacusprofundi ATCC 49239	Archaea	Euryarchaeota	Halobacteria	64.0	64.9	CP001365.1	2607869	2608900	1032
radA	Halosimplex carlsbadense 2-9-1	Archaea	Euryarchaeota	Halobacteria	67.7	67.3	NZ_AOIU01000018.1	120432	121478	1047
radA	Halostagnicola larsenii XH-48	Archaea	Euryarchaeota	Halobacteria	60.9	62.6	CP007055.1	78213	79244	1032
radA	Haloterrigena turkmenica DSM 5511	Archaea	Euryarchaeota	Halobacteria	64.1	65.6	CP001860.1	231614	232645	1032
radA	Halovivax ruber XH-70	Archaea	Euryarchaeota	Halobacteria	64.3	64.1	CP003050.1	2782573	2783607	1035
radA	Natrialba magadii ATCC 43099	Archaea	Euryarchaeota	Halobacteria	61.1	62.0	CP001932.1	1856981	1858012	1032
radA	Natrinema pellirubrum DSM 15624	Archaea	Euryarchaeota	Halobacteria	64.0	65.2	CP003372.1	3677598	3678629	1032
radA	Natronobacterium gregoryi SP2	Archaea	Euryarchaeota	Halobacteria	62.3	64.1	CP003377.1	1753916	1754947	1032
radA	Natronococcus occultus SP4	Archaea	Euryarchaeota	Halobacteria	64.6	64.6	CP003929.1	91533	92564	1032
radA	Salinararchaeum sp. Harcht-Bsk1	Archaea	Euryarchaeota	Halobacteria	66.6	67.2	CP005962.1	1726915	1727949	1035
radA	Methanobacterium sp. MB1	Archaea	Euryarchaeota	Methanobacteria	39.2	46.0	HG425166.1	1898524	1899459	936
radA	Methanobrevibacter smithii ATCC 35061	Archaea	Euryarchaeota	Methanobacteria	30.5	34.2	CP000678.1	1355135	1356070	936

radA	Methanosphaera stadtmanae DSM 3091	Archaea	Euryarchaeota	Methanobacteria	27.6	33.8	CP000102.1	1671474	1672409	936
radA	Methanothermobacter thermoautotrophicus str. Delta H	Archaea	Euryarchaeota	Methanobacteria	49.5	52.5	AE000666.1	1250239	1251174	936
radA	Methanothermobacter fervidus DSM 2088	Archaea	Euryarchaeota	Methanobacteria	31.6	35.9	CP002278.1	181701	182636	936
radA	Methanocaldococcus jannaschii DSM 2661	Archaea	Euryarchaeota	Methanococci;	31.3	36.3	L77117.1	791860	792918	1059
radA	Methanococcus maripaludis C7	Archaea	Euryarchaeota	Methanococci;	33.1	37.9	CP000745.1	480003	480971	969
radA	Methanothermococcus okinawensis IH1	Archaea	Euryarchaeota	Methanococci;	29.3	37.2	CP002792.1	220812	221780	969
radA	Methanoterris igneus Kol.5	Archaea	Euryarchaeota	Methanococci;	32.3	37.4	CP002737.1	255	1223	969
radA	Candidatus Methanoregula boonei 6A8	Archaea	Euryarchaeota	Methanomicrobia	54.5	55.5	CP000780.1	2173052	2174029	978
radA	Candidatus Methanosphaerula palustris E1-9c	Archaea	Euryarchaeota	Methanomicrobia	55.4	60.1	CP001338.1	39098	40081	984
radA	Methanocella arvonyzae MRE50	Archaea	Euryarchaeota	Methanomicrobia	54.6	59.2	AM114193.2	1896230	1897201	972
radA	Methanococcoides burtonii DSM 6242	Archaea	Euryarchaeota	Methanomicrobia	40.8	51.1	CP000300.1	2138696	2139673	978
radA	Methanocorpusculum labreanum Z	Archaea	Euryarchaeota	Methanomicrobia	50.0	56.7	CP000559.1	612493	613482	990
radA	Methanoculleus marisnigri JR1	Archaea	Euryarchaeota	Methanomicrobia	62.1	64.0	CP000562.1	1913934	1914908	975
radA	Methanohalobium evestigatum Z-7303	Archaea	Euryarchaeota	Methanomicrobia	36.4	43.3	CP002069.1	2166691	2167668	978
radA	Methanohalophilus mahii DSM 5219	Archaea	Euryarchaeota	Methanomicrobia	42.6	47.8	CP001994.1	273207	274184	978
radA	Methanobolus psychrophilus R15	Archaea	Euryarchaeota	Methanomicrobia	44.6	50.1	CP003083.1	1944792	1945769	978
radA	Methanomethylovorans hollandica DSM 15978	Archaea	Euryarchaeota	Methanomicrobia	41.8	48.9	CP003362.1	2252574	2253551	978
radA	Methanoplanus petrolearius DSM 11571	Archaea	Euryarchaeota	Methanomicrobia	47.4	53.4	CP002117.1	720935	721906	972
radA	Methanosaepta thermophila PT	Archaea	Euryarchaeota	Methanomicrobia	53.5	60.0	CP000477.1	959678	960646	969
radA	Methanosalsum zhilinae DSM 4017	Archaea	Euryarchaeota	Methanomicrobia	39.2	43.8	CP002101.1	2035075	2036052	978
radA	Methanosarcina barkeri str. Fusaro	Archaea	Euryarchaeota	Methanomicrobia	39.2	49.2	CP000099.1	2963521	2964498	978
radA	Methanospirillum hungatei JF-1	Archaea	Euryarchaeota	Methanomicrobia	45.1	52.5	CP000254.1	3213332	3214555	1224
radA	Methanopyrus kandleri AV19	Archaea	Euryarchaeota	Methanopyri	61.2	59.7	AE009439.1	1474919	1475872	954
radA	Palaeococcus pacificus DY20341	Archaea	Euryarchaeota	Thermococci	43.0	43.7	CP006019.1	1856108	1857178	1071
radA	Pyrococcus furiosus DSM 3638	Archaea	Euryarchaeota	Thermococci	40.8	43.8	AE009950.1	1776415	1777464	1050
radA	Thermococcus onnurineus NA1	Archaea	Euryarchaeota	Thermococci	51.3	54.5	CP000855.1	1512120	1513181	1062
radA	Candidatus Methanomethylophilus alvus Mx1201	Archaea	Euryarchaeota	Thermoplasmata	55.6	57.8	CP004049.1	228154	229104	951
radA	Ferroplasma acidarmanus fer1	Archaea	Euryarchaeota	Thermoplasmata	36.5	42.4	CP004145.1	1533509	1534504	996
radA	Picrophilus torridus DSM 9790	Archaea	Euryarchaeota	Thermoplasmata	36.0	44.7	AE017261.1	756544	757515	972
radA	Thermoplasma volcanium GSS1	Archaea	Euryarchaeota	Thermoplasmata	39.9	46.0	BA000011.4	443339	444310	972
radA	Thermoplasmatales archaeon BRNA1	Archaea	Euryarchaeota	Thermoplasmata	58.3	61.8	CP002916.1	244020	244964	945
radA	Candidatus Korarchaeum cryptofilum OPF8	Archaea	TACK group	Thermoplasmata	49.0	50.0	CP000968.1	1219587	1220597	1011
radA	Acidianus hospitalis W1	Archaea	TACK group	Crenarchaeota	34.1	38.7	CP002535.1	724793	725710	918
radA	Aeropyrum pernix K1	Archaea	TACK group	Crenarchaeota	56.3	53.2	BA000002.3	88998	89957	960
radA	Caldisphaera lagunensis DSM 15908	Archaea	TACK group	Crenarchaeota	30.0	36.2	CP003378.1	4945	5928	984
radA	Caldivirga maquilingensis IC-167	Archaea	TACK group	Crenarchaeota	43.1	47.2	CP000852.1	1174551	1175627	1077
radA	Desulfurococcus kamchatkensis 1221n	Archaea	TACK group	Crenarchaeota	45.3	47.9	CP001140.1	1136697	1137683	987
radA	Fervidococcus fontis Kam940	Archaea	TACK group	Crenarchaeota	37.5	39.0	CP003423.1	715696	717243	1548
radA	Hyperthermus butylicus DSM 5456	Archaea	TACK group	Crenarchaeota	53.7	50.3	CP000493.1	979368	980339	972
radA	Ignicoccus hospitalis KIN4/I	Archaea	TACK group	Crenarchaeota	56.5	58.5	CP000816.1	896102	897085	984
radA	Ignisphaera aggregans DSM 17230	Archaea	TACK group	Crenarchaeota	35.7	38.4	CP002098.1	1358182	1359132	951
radA	Metallosphaera sedula DSM 5348	Archaea	TACK group	Crenarchaeota	46.2	47.9	CP000682.1	6305	7279	975
radA	Pyrobaculum aerophilum str. IM2	Archaea	TACK group	Crenarchaeota	51.4	50.9	AE009441.1	351906	352907	1002
radA	Pyrodicticum delaneyi strain Su06	Archaea	TACK group	Crenarchaeota	53.9	48.0	CP013011.1	1709137	1710108	972
radA	Pyrolobus fumarii 1A	Archaea	TACK group	Crenarchaeota	54.9	54.0	CP002838.1	920225	921193	969
radA	Staphylothermus marinus F1	Archaea	TACK group	Crenarchaeota	35.7	40.9	CP000575.1	952808	953767	960
radA	Sulfolobus islandicus L.S.2.15	Archaea	TACK group	Crenarchaeota	35.1	36.8	CP001399.1	1820289	1821263	975
radA	Thermofilium sp. 1807-2	Archaea	TACK group	Crenarchaeota	47.2	50.4	CP009961.1	920272	921264	993
radA	Thermogladius cellulolyticus 1633	Archaea	TACK group	Crenarchaeota	55.6	56.8	CP003531.1	225126	226094	969
radA	Thermoproteus uzoniensis 768-20	Archaea	TACK group	Crenarchaeota	59.7	58.9	CP002590.1	1579386	1580399	1014
radA	Thermosphaera aggregans DSM 11486	Archaea	TACK group	Crenarchaeota	46.7	56.6	CP001939.1	718116	719096	981

radA	Vulcanisaeta distributa DSM 14429	572478	Archaea	TACK group	Crenarchaeota	45.7	49.3	CP002100.1	2346532	2347698	1167
radA	Nitrosopumilus maritimus SCM1	436308	Archaea	TACK group	Thaumarchaeota	34.2	39.6	CP000866.1	1268792	1269958	1167
radA	Nitrososphaera viennensis EN76	926571	Archaea	TACK group	Thaumarchaeota	52.7	55.3	CP007536.1	440254	441273	1020
radA	Thaumarchaeota archaeon MY3	1353260	Archaea	TACK group	Thaumarchaeota	34.1	38.0	CP012850.1	752802	753914	1113
radA	Archaeon GW2011_AR10	1579370	Archaea	unclassified Archaea	unclassified Archaea	43.1	43.8	CP010424.1	1024229	1025176	948
radA	Halophilic archaeon DL31	756883	Archaea	unclassified Archaea	unclassified Archaea	62.4	64.0	CP002988.1	2564138	2565169	1032
recA	Amycolatopsis mediterranei U32	749927	Bacteria	Actinobacteria	Actinobacteria	71.3	68.2	NC_014318.1	2502300	2503346	1047
recA	Bifidobacterium bifidum PRL2010	702459	Bacteria	Actinobacteria	Actinobacteria	62.7	63.2	NC_014638.1	1489100	1490290	1191
recA	Bifidobacterium longum NCC2705	206672	Bacteria	Actinobacteria	Actinobacteria	59.9	62.4	NC_004307.2	326075	327268	1191
recA	Corynebacterium glutamicum ATCC 13032	196627	Bacteria	Actinobacteria	Actinobacteria	54.0	53.4	NC_003450.3	2063986	2065116	1134
recA	Gardnerella vaginalis ATCC 14019	525284	Bacteria	Actinobacteria	Actinobacteria	42.0	40.0	NC_014644.1	1059017	1060171	1155
recA	Glaciibacter superstes DSM 21135	1121924	Bacteria	Actinobacteria	Actinobacteria	64.4	63.0	NZ_ATWH01000015.1	80176	81279	1104
recA	Mycobacterium leprae TN	272631	Bacteria	Actinobacteria	Actinobacteria	57.8	56.0	NC_002677.1	1153056	1155191	2136
recA	Mycobacterium smegmatis str. MC2 155	246196	Bacteria	Actinobacteria	Actinobacteria	67.4	66.0	NC_008596.1	2792363	2793412	1050
recA	Mycobacterium tuberculosis H37Rv	83332	Bacteria	Actinobacteria	Actinobacteria	65.6	63.3	NC_000962.3	3049052	3051424	2373
recA	Streptomyces coelicolor A3(2)	100226	Bacteria	Actinobacteria	Actinobacteria	72.0	68.4	NC_003888.3	6306654	6307778	1125
recA	Aquifex aeolicus VF5	224324	Bacteria	Aquificae	Aquificae	43.3	44.0	NC_000918.1	1515870	1516907	1038
recA	Bacteroides thetaiotaomicron VPI-5482	226186	Bacteria	Bacteroidetes/Chlorobi group	Bacteroidetes	43.0	43.3	NC_004663.1	6035657	6036691	1035
recA	Flavobacterium psychrophilum JIP02/86	402612	Bacteria	Bacteroidetes/Chlorobi group	Bacteroidetes	32.5	38.1	NC_009613.3	2582267	2583271	1005
recA	Salinibacter ruber DSM 13855	309807	Bacteria	Bacteroidetes/Chlorobi group	Bacteroidetes	66.0	62.9	NC_007677.1	1907949	1909010	1062
recA	Chlorobium tepidum TLS	194439	Bacteria	Bacteroidetes/Chlorobi group	Chlorobi	56.5	60.7	NC_002932.3	1824539	1825579	1041
recA	Chlamydia psittaci	83554	Bacteria	Chlamydiae/Verrucomicrobia group	Chlamydiae	39.1	40.5	NC_017287.1	1154275	1155324	1050
recA	Chlamydia trachomatis 434/Bu	471472	Bacteria	Chlamydiae/Verrucomicrobia group	Chlamydiae	41.3	44.7	NC_010287.1	25080	26138	1059
recA	Chlamydia trachomatis DUW-3/CX	272561	Bacteria	Chlamydiae/Verrucomicrobia group	Chlamydiae	41.3	44.1	NC_000117.1	745340	746398	1059
recA	Chlamydomydia pneumoniae CWL029	115713	Bacteria	Chlamydiae/Verrucomicrobia group	Chlamydiae	40.6	41.7	NC_000922.1	858536	859597	1062
recA	Chloroflexus aurantiacus J-10-fl	324602	Bacteria	Chloroflexi	Chloroflexia	56.7	56.8	NC_010175.1	1875019	1876074	1056
recA	Gloeobacter violaceus PCC 7421	251221	Bacteria	Cyanobacteria	Gloeobacteria	62.0	63.7	NC_005125.1	3612670	3613779	1110
recA	Nostoc azollae 0708	551115	Bacteria	Cyanobacteria	Nostocales	38.4	43.4	NC_014248.1	756110	757189	1080
recA	Nostoc punctiforme ATCC 29133	63737	Bacteria	Cyanobacteria	Nostocales	41.4	44.8	NC_010628.1	3601438	3602517	1080
recA	Acarochloris marina MBIC11017	329726	Bacteria	Cyanobacteria	Oscillatoriothycideae	47.3	51.7	NC_009925.1	5078111	5079169	1059
recA	Arthrospira platensis NIES-39	696747	Bacteria	Cyanobacteria	Oscillatoriothycideae	44.3	47.4	U39924.1	54	1151	1098
recA	Cyanothece sp. ATCC 51142	43989	Bacteria	Cyanobacteria	Oscillatoriothycideae	39.2	44.6	CP000806.1	4791407	4792456	1050
recA	Microcystis aeruginosa NIES-843	449447	Bacteria	Cyanobacteria	Oscillatoriothycideae	42.6	48.6	NC_010296.1	1445173	1446240	1068
recA	Oscillatoria sp. PCC 7112	179408	Bacteria	Cyanobacteria	Oscillatoriothycideae	45.8	48.9	NC_019729.1	831814	832953	1140
recA	Synechococcus elongatus PCC 7942	1140	Bacteria	Cyanobacteria	Oscillatoriothycideae	55.5	58.4	NC_007604.1	340698	341768	1071
recA	Synechocystis sp. PCC 6803	1148	Bacteria	Cyanobacteria	Oscillatoriothycideae	47.4	53.0	NC_000911.1	3549956	3551020	1065
recA	Thermosynechococcus elongatus BP-1	197221	Bacteria	Cyanobacteria	Oscillatoriothycideae	53.9	52.9	NC_004113.1	2171969	2173036	1068
recA	Trichodesmium erythraeum IMS101	203124	Bacteria	Cyanobacteria	Oscillatoriothycideae	35.7	39.7	NC_008312.1	4169525	4170607	1083
recA	Prochlorococcus marinus subsp. marinus str. CCMP1375	167539	Bacteria	Cyanobacteria	Prochlorales	31.2	42.1	NC_005042.1	1567932	1569071	1140
recA	Prochlorococcus marinus subsp. pastoris str. CCMP1986	59919	Bacteria	Cyanobacteria	Prochlorales	30.8	35.9	NC_005072.1	1490458	1491570	1113
recA	Deinococcus radiodurans R1	243230	Bacteria	Deinococcus-Thermus	Deinococci	66.8	66.1	NC_001263.1	2337795	2338886	1092
recA	Thermus thermophilus HB8	300852	Bacteria	Deinococcus-Thermus	Deinococci	69.4	70.3	NC_006461.1	1701658	1702680	1023
recA	Bacillus anthracis str. Ames	198094	Bacteria	Firmicutes	Bacilli	35.2	40.5	NC_003997.3	3590268	3591626	1359
recA	Bacillus cereus ATCC 14579	226900	Bacteria	Firmicutes	Bacilli	35.2	40.9	NC_004722.1	3751511	3752542	1032
recA	Bacillus subtilis subsp. subtilis str. 168	224308	Bacteria	Firmicutes	Bacilli	43.7	46.7	NC_000964.3	1764645	1765691	1047
recA	Enterococcus faecalis V583	226185	Bacteria	Firmicutes	Bacilli	37.3	42.6	NC_004668.1	3044471	3045517	1047
recA	Enterococcus faecium DO	333849	Bacteria	Firmicutes	Bacilli	38.0	41.9	NC_017960.1	2571787	2572836	1050
recA	Lactobacillus acidophilus NCFM	272621	Bacteria	Firmicutes	Bacilli	34.7	39.7	NC_006814.3	651885	652976	1092
recA	Lactobacillus plantarum WCFS1	220668	Bacteria	Firmicutes	Bacilli	44.5	47.6	NC_004567.2	2078955	2080097	1143
recA	Lactobacillus salivarius UCC118	362948	Bacteria	Firmicutes	Bacilli	32.9	37.0	NC_007929.1	1161704	1162849	1146
recA	Lactococcus lactis subsp. lactis Il1403	272623	Bacteria	Firmicutes	Bacilli	35.2	39.7	NC_002662.1	358813	359976	1164

recA	Listeria monocytogenes EGD-e	169963	Bacteria	Firmicutes	Bacilli	37.9	39.3	NC_003210.1	1425419	1426465	1047
recA	Lysinibacillus sphaericus C3-41	444177	Bacteria	Firmicutes	Bacilli	37.3	39.6	CP000817.1	1719130	1720119	990
recA	Staphylococcus aureus subsp. aureus NCTC 8325	93061	Bacteria	Firmicutes	Bacilli	32.8	37.0	NC_007795.1	1216223	1217266	1044
recA	Staphylococcus epidermidis ATCC 12228	176280	Bacteria	Firmicutes	Bacilli	32.0	36.7	NC_004461.1	969196	970245	1050
recA	Streptococcus agalactiae 2603V/R	208435	Bacteria	Firmicutes	Bacilli	35.3	39.6	NC_004116.1	2072723	2073862	1140
recA	Streptococcus mitis B6	365659	Bacteria	Firmicutes	Bacilli	40.1	43.4	NC_013853.1	1901488	1902642	1155
recA	Streptococcus mutans UA159	210007	Bacteria	Firmicutes	Bacilli	36.8	39.1	NC_004350.2	1956459	1957610	1152
recA	Streptococcus pneumoniae R6	171101	Bacteria	Firmicutes	Bacilli	39.6	41.8	NC_003098.1	1725078	1726244	1167
recA	Streptococcus pyogenes M1 GAS	160490	Bacteria	Firmicutes	Bacilli	38.5	40.7	NC_002737.2	1766408	1767544	1137
recA	Streptococcus sanguinis SK36	388919	Bacteria	Firmicutes	Bacilli	43.2	45.7	NC_009009.1	2240051	2241199	1149
recA	Streptococcus suis BM407	568814	Bacteria	Firmicutes	Bacilli	41.2	43.9	NC_012926.1	67689	68840	1152
recA	Clostridium acetobutylicum ATCC 824	272562	Bacteria	Firmicutes	Clostridia	30.9	33.6	NC_003030.1	1965171	1966220	1050
recA	Clostridium botulinum A str. Hall	441771	Bacteria	Firmicutes	Clostridia	28.0	32.6	NC_009698.1	2376995	2378053	1059
recA	Clostridium tetani 12124569	1231072	Bacteria	Firmicutes	Clostridia	28.5	33.9	NC_022777.1	1360630	1361610	981
recA	Peptoclostridium difficile 630	272563	Bacteria	Firmicutes	Clostridia	28.4	33.3	NC_009089.1	1539910	1540956	1047
recA	Fusobacterium nucleatum subsp. nucleatum ATCC 25586	190304	Bacteria	Fusobacteria	Fusobacteria	27.0	33.2	NC_003454.1	1199783	1200928	1146
recA	Thermodesulfobivrio yellowstonii DSM 11347	289376	Bacteria	Nitrospirae	Nitrospira	34.1	39.0	NC_011296.1	519949	520977	1029
recA	Rhodopirellula baltica SH 1	243090	Bacteria	Planctomycetes	Planctomycetia	55.5	56.8	NC_005027.1	6387411	6388538	1128
recA	Agrobacterium fabrum str. C58	176299	Bacteria	Proteobacteria	Alphaproteobacteria	59.1	59.2	NC_003062.2	1853926	1855017	1092
recA	Bradyrhizobium diazoefficiens USDA 110	224911	Bacteria	Proteobacteria	Alphaproteobacteria	64.1	64.4	NC_004463.1	6324669	6325757	1089
recA	Caulobacter crescentus CB15	190650	Bacteria	Proteobacteria	Alphaproteobacteria	67.2	64.1	NC_002696.2	1219186	1220256	1071
recA	Ketogulonicigenium vulgare WSH-001	759362	Bacteria	Proteobacteria	Alphaproteobacteria	61.7	59.0	NC_017384.1	1013652	1014740	1089
recA	Neorickettsia risticii	950	Bacteria	Proteobacteria	Alphaproteobacteria	41.3	43.8	NC_013009.1	468288	469271	984
recA	Rhodococcus sphaeroides 2.4.1	272943	Bacteria	Proteobacteria	Alphaproteobacteria	68.8	63.7	NC_007493.2	2184378	2185451	1074
recA	Rhodospirillum rubrum ATCC 11170	269796	Bacteria	Proteobacteria	Alphaproteobacteria	65.4	63.6	NC_007643.1	3252155	3253255	1101
recA	Rickettsia canadensis	788	Bacteria	Proteobacteria	Alphaproteobacteria	31.1	37.2	NC_009879.1	990368	991399	1032
recA	Rickettsia prowazekii str. Madrid E	272947	Bacteria	Proteobacteria	Alphaproteobacteria	29.0	36.2	NC_000963.1	954073	955095	1023
recA	Rickettsia rickettsii str. Iowa	452659	Bacteria	Proteobacteria	Alphaproteobacteria	32.5	37.9	NC_010263.2	1084779	1085867	1089
recA	Sinorhizobium fredii NGR234	394	Bacteria	Proteobacteria	Alphaproteobacteria	62.2	60.2	NC_012587.1	1694044	1695129	1086
recA	Sinorhizobium medicae WSM419	366394	Bacteria	Proteobacteria	Alphaproteobacteria	61.2	59.0	NC_009636.1	1606902	1607987	1086
recA	Sinorhizobium meliloti 1021	266834	Bacteria	Proteobacteria	Alphaproteobacteria	62.0	59.9	NC_003047.1	1948178	1949263	1086
recA	Achromobacter xylosoxidans A8	762376	Bacteria	Proteobacteria	Betaproteobacteria	67.7	63.2	NC_014640.1	1985704	1986768	1065
recA	Bordetella bronchiseptica 253	568707	Bacteria	Proteobacteria	Betaproteobacteria	68.2	64.2	NC_019382.1	1535561	1536622	1062
recA	Bordetella hinzii OH87 BAL007II	1331262	Bacteria	Proteobacteria	Betaproteobacteria	67.0	63.9	NZ_JHEM01000010.1	292468	293529	1062
recA	Bordetella holmesii ATCC 51541	1247649	Bacteria	Proteobacteria	Betaproteobacteria	62.7	58.4	NZ_CP007494.1	2945457	2946518	1062
recA	Bordetella parapertussis Bpp5	1208660	Bacteria	Proteobacteria	Betaproteobacteria	67.9	64.3	NC_018828.1	2137904	2138965	1062
recA	Bordetella pertussis Tohama I	257313	Bacteria	Proteobacteria	Betaproteobacteria	67.7	64.3	NC_002929.2	2698234	2699295	1062
recA	Burkholderia cenocepacia J2315	216591	Bacteria	Proteobacteria	Betaproteobacteria	66.9	63.6	NC_011000.1	1041142	1042212	1071
recA	Burkholderia glumae BGR1	626418	Bacteria	Proteobacteria	Betaproteobacteria	68.2	64.9	NC_012724.2	3392994	3394076	1083
recA	Burkholderia mallei ATCC 23344	243160	Bacteria	Proteobacteria	Betaproteobacteria	68.5	64.1	NC_006348.1	290127	291197	1071
recA	Burkholderia multivorans	87883	Bacteria	Proteobacteria	Betaproteobacteria	66.9	64.6	NZ_CP008730.1	2491085	2492158	1074
recA	Burkholderia pseudomallei K96243	272560	Bacteria	Proteobacteria	Betaproteobacteria	68.2	64.1	NC_006350.1	903338	904408	1071
recA	Burkholderia xenovorans LB400	266265	Bacteria	Proteobacteria	Betaproteobacteria	62.6	59.1	NC_007951.1	4141625	4142701	1077
recA	Burkholderia zhejiangensis	871203	Bacteria	Proteobacteria	Betaproteobacteria	62.8	62.3	NZ_JFHD01000002.1	483316	484386	1071
recA	Comamonas testosteroni CNB-2	688245	Bacteria	Proteobacteria	Betaproteobacteria	61.4	61.6	NC_013446.2	5346122	5347231	1110
recA	Cupriavidus metallidurans CH34	266264	Bacteria	Proteobacteria	Betaproteobacteria	63.5	63.7	NC_007973.1	495632	496693	1062
recA	Defluviacidovorans SPH-1	398578	Bacteria	Proteobacteria	Betaproteobacteria	66.5	64.8	NC_010002.1	980943	982046	1104
recA	Herbaspirillum seropedicae SmR1	757424	Bacteria	Proteobacteria	Betaproteobacteria	63.4	62.3	NC_014323.1	719977	721032	1056
recA	Kingella kingae ATCC 23330	887327	Bacteria	Proteobacteria	Betaproteobacteria	46.7	48.4	NZ_GL891959.1	202104	203213	1110
recA	Neisseria gonorrhoeae FA 1090	242231	Bacteria	Proteobacteria	Betaproteobacteria	52.4	56.2	NC_002946.2	737184	738230	1047
recA	Neisseria meningitidis MC58	122586	Bacteria	Proteobacteria	Betaproteobacteria	51.9	56.3	NC_003112.2	1488548	1489594	1047

recA	<i>Neisseria mucosa</i> C102	435832	Bacteria	Proteobacteria	Betaproteobacteria	49.7	52.0	NZ_GL635793.1	536717	537763	1047
recA	<i>Neisseria sicca</i> ATCC 29256	547045	Bacteria	Proteobacteria	Betaproteobacteria	51.1	56.4	NZ_ACKO02000010.1	58893	59939	1047
recA	<i>Pandoraea pnomenusa</i>	93220	Bacteria	Proteobacteria	Betaproteobacteria	64.7	62.2	NZ_CP009553.1	2335977	2337038	1062
recA	<i>Ralstonia eutropha</i> H16	381666	Bacteria	Proteobacteria	Betaproteobacteria	65.5	65.1	NC_008313.1	571475	572539	1065
recA	<i>Ralstonia pickettii</i> 12J	402626	Bacteria	Proteobacteria	Betaproteobacteria	63.6	61.8	NC_010682.1	504956	506014	1059
recA	<i>Ralstonia solanacearum</i> GM11000	267608	Bacteria	Proteobacteria	Betaproteobacteria	66.5	64.8	NC_003295.1	596177	597235	1059
recA	<i>Campylobacter jejuni</i> subsp. jejuni NCTC 11168 ATCC 700819	192222	Bacteria	Proteobacteria	delta/epsilon subdivisions	30.4	36.2	NC_002163.1	1593765	1594796	1032
recA	<i>Desulfovibrio vulgaris</i> str. Hildenborough	882	Bacteria	Proteobacteria	delta/epsilon subdivisions	63.2	63.1	NC_002937.3	1194497	1195570	1074
recA	<i>Geobacter sulfurreducens</i> PCA	243231	Bacteria	Proteobacteria	delta/epsilon subdivisions	60.9	62.9	NC_002939.5	161048	162064	1017
recA	<i>Helicobacter pylori</i> 26695	85962	Bacteria	Proteobacteria	delta/epsilon subdivisions	38.9	44.3	NC_000915.1	162928	163971	1044
recA	<i>Acinetobacter calcoaceticus</i> PHEA-2	871585	Bacteria	Proteobacteria	Gammaproteobacteria	38.7	41.2	NC_016603.1	1550935	1551981	1047
recA	<i>Aeromonas hydrophila</i> subsp. hydrophila ATCC 7966	380703	Bacteria	Proteobacteria	Gammaproteobacteria	61.5	59.8	NC_008570.1	4147646	4148710	1065
recA	<i>Candidatus Carsonella ruddii</i> HT isolate Thao2000	1202539	Bacteria	Proteobacteria	Gammaproteobacteria	14.6	16.6	NC_018417.1	100513	101403	891
recA	<i>Citrobacter freundii</i> CFNIH1	1333848	Bacteria	Proteobacteria	Gammaproteobacteria	51.5	53.4	NZ_CP007557.1	100209	101273	1065
recA	<i>Coxiella burnetii</i> RSA 493	227377	Bacteria	Proteobacteria	Gammaproteobacteria	42.6	47.0	NC_002971.3	997796	998827	1032
recA	<i>Cronobacter sakazakii</i> ATCC BAA-894	290339	Bacteria	Proteobacteria	Gammaproteobacteria	56.8	55.7	NC_009778.1	540596	541660	1065
recA	<i>Dickeya dadantii</i> 3937	198628	Bacteria	Proteobacteria	Gammaproteobacteria	56.3	54.0	NC_014500.1	3702951	3704027	1077
recA	<i>Dickeya solani</i> IPO 2222	1225786	Bacteria	Proteobacteria	Gammaproteobacteria	56.2	52.9	NZ_CM001859.1	3616431	3617507	1077
recA	<i>Enterobacter aerogenes</i> KCTC 2190	1028307	Bacteria	Proteobacteria	Gammaproteobacteria	55.0	56.9	NC_015663.1	316232	317290	1059
recA	<i>Enterobacter lignolyticus</i> SCF1	701347	Bacteria	Proteobacteria	Gammaproteobacteria	57.0	56.5	NC_014618.1	1075207	1076265	1059
recA	<i>Erwinia amylovora</i> CFBP1430	665029	Bacteria	Proteobacteria	Gammaproteobacteria	53.6	53.2	NC_013961.1	889989	891056	1068
recA	<i>Escherichia coli</i> IA139	585057	Bacteria	Proteobacteria	Gammaproteobacteria	50.6	53.9	NC_011750.1	2993159	2994220	1062
recA	<i>Escherichia coli</i> O104:H4 str. 2011C-3493	1133852	Bacteria	Proteobacteria	Gammaproteobacteria	50.6	53.9	NC_018658.1	1248409	1249470	1062
recA	<i>Escherichia coli</i> O157:H7 str. Sakai	386585	Bacteria	Proteobacteria	Gammaproteobacteria	50.6	53.9	NC_002695.1	3546635	3547696	1062
recA	<i>Escherichia coli</i> O83:H1 str. NRG 857C	685038	Bacteria	Proteobacteria	Gammaproteobacteria	50.6	53.8	NC_017634.1	2779679	2780740	1062
recA	<i>Escherichia coli</i> str. K-12 substr. MG1655	511145	Bacteria	Proteobacteria	Gammaproteobacteria	50.6	54.1	NC_000913.3	2822708	2823769	1062
recA	<i>Escherichia coli</i> UMN026	585056	Bacteria	Proteobacteria	Gammaproteobacteria	50.6	53.7	NC_011751.1	3128409	3129470	1062
recA	<i>Francisella tularensis</i> subsp. tularensis SCHU S4	177416	Bacteria	Proteobacteria	Gammaproteobacteria	32.3	38.1	NC_006570.2	1838207	1839286	1080
recA	<i>Haemophilus influenzae</i> Rd KW20	71421	Bacteria	Proteobacteria	Gammaproteobacteria	38.0	39.2	NC_000907.1	621492	622556	1065
recA	<i>Kosakonia sacchari</i> SP1	1235834	Bacteria	Proteobacteria	Gammaproteobacteria	53.7	53.7	NZ_CP007215.1	4222223	4223287	1065
recA	<i>Legionella pneumophila</i> subsp. pneumophila str. Philadelphia 1	272624	Bacteria	Proteobacteria	Gammaproteobacteria	38.2	41.1	NC_002942.5	2010421	2011467	1047
recA	<i>Methylobacterium alcaliphilum</i> 20Z	1091494	Bacteria	Proteobacteria	Gammaproteobacteria	48.7	50.1	NC_016112.1	3261633	3262673	1041
recA	<i>Morganella morganii</i> subsp. morganii KT	1124991	Bacteria	Proteobacteria	Gammaproteobacteria	50.9	50.3	NC_020418.1	1467473	1468543	1071
recA	<i>Pantoea ananatis</i> LMG 5342	1123863	Bacteria	Proteobacteria	Gammaproteobacteria	53.6	51.6	NC_016816.1	1028486	1029556	1071
recA	<i>Pectobacterium atrosepticum</i> SCRI1043	218491	Bacteria	Proteobacteria	Gammaproteobacteria	51.0	48.4	NC_004547.2	3785644	3786717	1074
recA	<i>Photobacterium leiognathi</i> subsp. mandapamensis svers. 1.1.	1001530	Bacteria	Proteobacteria	Gammaproteobacteria	41.2	44.5	NZ_DF093597.1	239373	240425	1053
recA	<i>Proteus mirabilis</i> HI4320	529507	Bacteria	Proteobacteria	Gammaproteobacteria	38.7	40.8	NC_010554.1	425599	426666	1068
recA	<i>Providencia rettgeri</i> Dmel1	1141663	Bacteria	Proteobacteria	Gammaproteobacteria	40.4	40.2	NZ_CM001774.1	3749063	3750130	1068
recA	<i>Pseudomonas aeruginosa</i> PAO1	208964	Bacteria	Proteobacteria	Gammaproteobacteria	66.3	64.2	NC_002516.2	4051564	4052604	1041
recA	<i>Pseudomonas mendocina</i> ymp	399739	Bacteria	Proteobacteria	Gammaproteobacteria	63.9	51.7	NC_009439.1	449065	450111	1047
recA	<i>Pseudomonas putida</i> KT2440	160488	Bacteria	Proteobacteria	Gammaproteobacteria	61.8	59.3	NC_002947.3	1827143	1828210	1068
recA	<i>Pseudomonas syringae</i> pv. <i>syringae</i> B728a	205918	Bacteria	Proteobacteria	Gammaproteobacteria	58.8	58.6	NC_007005.1	1562480	1563544	1065
recA	<i>Pseudomonas syringae</i> pv. <i>tomato</i> str. DC3000	223283	Bacteria	Proteobacteria	Gammaproteobacteria	58.6	58.4	NC_004578.1	4540716	4541780	1065
recA	<i>Salmonella enterica</i> subsp. <i>enterica</i> serovar <i>Typhi</i> str. CT18	220341	Bacteria	Proteobacteria	Gammaproteobacteria	52.1	54.4	NC_003198.1	2827753	2828814	1062
recA	<i>Salmonella enterica</i> subsp. <i>enterica</i> serovar <i>Typhimurium</i> str. LT2	99287	Bacteria	Proteobacteria	Gammaproteobacteria	52.1	54.7	NC_003197.1	2974854	2975915	1062
recA	<i>Serratia marcescens</i> subsp. <i>marcescens</i> Db11	273526	Bacteria	Proteobacteria	Gammaproteobacteria	59.5	57.9	NZ_HG326223.1	179686	180750	1065
recA	<i>Shewanella baltica</i> OS678	693973	Bacteria	Proteobacteria	Gammaproteobacteria	46.3	45.0	NC_016901.1	3796983	3798050	1068
recA	<i>Shewanella oneidensis</i> MR-1	211586	Bacteria	Proteobacteria	Gammaproteobacteria	45.9	46.1	NC_004347.2	3573159	3574232	1074
recA	<i>Shigella dysenteriae</i> Sd197	300267	Bacteria	Proteobacteria	Gammaproteobacteria	50.9	54.0	NC_007606.1	2676654	2677715	1062
recA	<i>Shigella flexneri</i> 2a str. 301	198214	Bacteria	Proteobacteria	Gammaproteobacteria	50.5	53.7	NC_004337.2	2796381	2797442	1062
recA	<i>Vibrio anguillarum</i> 775	882102	Bacteria	Proteobacteria	Gammaproteobacteria	44.4	44.2	NC_015633.1	2753533	2754579	1047

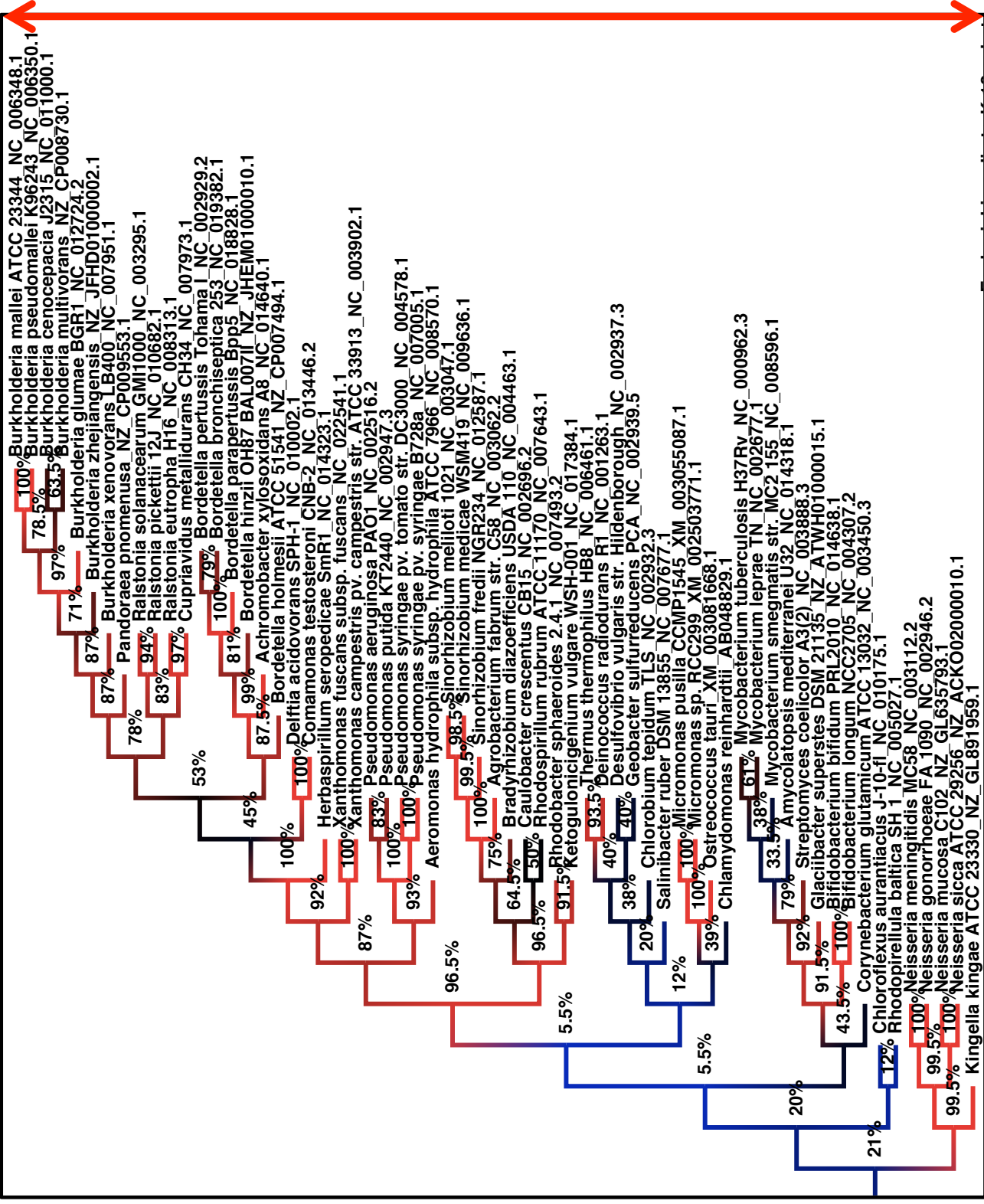
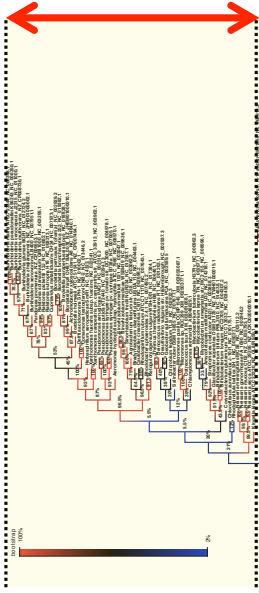
recA	<i>Vibrio campbellii</i> ATCC BAA-1116	338187	Bacteria	Proteobacteria	Gammaproteobacteria	45.5	46.3	NC_009783.1	3560849	3561892	1044
recA	<i>Vibrio cholerae</i> O1 biovar El Tor str. N16961	243277	Bacteria	Proteobacteria	Gammaproteobacteria	47.5	47.6	NC_002505.1	574696	575760	1065
recA	<i>Vibrio cyclitrophicus</i> FF75	1136163	Bacteria	Proteobacteria	Gammaproteobacteria	43.5	44.2	NZ_ATLT01000001.1	85683	86723	1041
recA	<i>Vibrio fischeri</i> ES114	312309	Bacteria	Proteobacteria	Gammaproteobacteria	38.3	40.7	NC_006840.2	577375	578421	1047
recA	<i>Vibrio nigripulchritudo</i>	28173	Bacteria	Proteobacteria	Gammaproteobacteria	45.8	47.5	NC_022528.1	144520	145560	1041
recA	<i>Vibrio orientalis</i> CIP 102891 ATCC 33934	675816	Bacteria	Proteobacteria	Gammaproteobacteria	44.8	46.6	NZ_ACZV01000005.1	2130292	2131398	1047
recA	<i>Vibrio parahaemolyticus</i> RIMD 2210633	223926	Bacteria	Proteobacteria	Gammaproteobacteria	45.3	46.5	NC_004603.1	2693309	2694352	1044
recA	<i>Vibrio tasmaniensis</i> LGP32	575788	Bacteria	Proteobacteria	Gammaproteobacteria	43.8	45.1	NC_011753.2	2783707	2784747	1041
recA	<i>Vibrio tubiashii</i> ATCC 19109	1051646	Bacteria	Proteobacteria	Gammaproteobacteria	45.0	47.4	NZ_CP009354.1	2836001	2837044	1044
recA	<i>Vibrio vulnificus</i> YJ016	196600	Bacteria	Proteobacteria	Gammaproteobacteria	46.7	47.0	NC_005139.1	2863509	2864558	1050
recA	<i>Xanthomonas campestris</i> pv. <i>campestris</i> str. ATCC 33913	190485	Bacteria	Proteobacteria	Gammaproteobacteria	65.3	63.7	NC_003902.1	2000077	2001108	1032
recA	<i>Xanthomonas fuscans</i> subsp. <i>fuscans</i>	366649	Bacteria	Proteobacteria	Gammaproteobacteria	64.8	61.5	NC_022541.1	3199814	3200848	1035
recA	<i>Xenorhabdus bovienii</i> SS-2004	406818	Bacteria	Proteobacteria	Gammaproteobacteria	44.7	44.8	NC_013892.1	3213002	3214078	1077
recA	<i>Xenorhabdus nematophila</i> ATCC 19061	406817	Bacteria	Proteobacteria	Gammaproteobacteria	43.7	44.9	NC_014228.1	1125811	1126887	1077
recA	<i>Yersinia enterocolitica</i> subsp. <i>enterocolitica</i> 8081	393305	Bacteria	Proteobacteria	Gammaproteobacteria	47.0	46.3	NC_008800.1	956095	957159	1065
recA	<i>Yersinia pestis</i> CO92	214092	Bacteria	Proteobacteria	Gammaproteobacteria	47.6	46.1	NC_003143.1	3687182	3688252	1071
recA	<i>Yersinia ruckeri</i> ATCC 29473	527005	Bacteria	Proteobacteria	Gammaproteobacteria	47.5	46.0	NZ_KN150747.1	1533142	1534212	1071
recA	<i>Borrelia burgdorferi</i> B31	224326	Bacteria	Spirochaetes	Spirochaetia	28.3	37.0	NC_001318.1	126316	127413	1098
recA	<i>Leptospira interrogans</i> serovar <i>Lai</i> str. 56601	189518	Bacteria	Spirochaetes	Spirochaetia	35.0	39.1	NC_004342.2	2150681	2151781	1101
recA	<i>Trponema denticola</i> ATCC 35405	243275	Bacteria	Spirochaetes	Spirochaetia	38.0	46.0	NC_002967.9	894981	896225	1245
recA	<i>Mycoplasma mycoides</i> subsp. <i>mycoides</i> SC str. PG1	272632	Bacteria	Tenericutes	Mollicutes	23.9	25.6	NC_005364.2	476358	477395	1038
recA	<i>Mycoplasma pneumoniae</i> M129	272634	Bacteria	Tenericutes	Mollicutes	40.0	45.5	NC_000912.1	596816	597826	1011
recA	<i>Thermotoga maritima</i> MSB8	243274	Bacteria	Thermotogae	Thermotogae	46.2	46.9	NC_000853.1	1842056	1843126	1071
recA	<i>Dictyostellium discoideum</i> AX4	352472	Eukaryota	Amoebozoa	Mycetozoa	22.5	28.8	XM_638224.1	1	1359	1359
recA	<i>Dictyostellium fasciculatum</i>	261658	Eukaryota	Amoebozoa	Mycetozoa	33.8	36.7	XM_004350318.1	1	1356	1356
recA	<i>Paulinella chromatophora</i>	39717	Eukaryota	Rhizaria	Cercozoa	38.0	40.8	CP000815.1	667656	668759	1104
recA	<i>Cyanidioschyzon merolae</i> strain 10D	280699	Eukaryota	Rhodophyta	Bangiophyceae	55.0	54.7	XM_005536433.1	1	1266	1266
recA	<i>Galdieria sulphuraria</i>	130081	Eukaryota	Rhodophyta	Bangiophyceae	37.9	42.9	XM_005704669.1	159	1349	1191
recA	<i>Phaeodactylum tricornutum</i> CCAP 1055/1	556484	Eukaryota	Stramenopiles	Bacillariophyta	48.8	51.0	XM_002177303.1	90	1385	1296
recA	<i>Chlamydomonas reinhardtii</i>	3055	Eukaryota	Viridiplantae	Chlorophyta	62.0	66.4	AB048829.1	121	1365	1245
recA	<i>Micromonas pusilla</i> CCMP1545	564608	Eukaryota	Viridiplantae	Chlorophyta	65.7	67.5	XM_003055087.1	1	1341	1341
recA	<i>Micromonas</i> sp. RCC299	296587	Eukaryota	Viridiplantae	Chlorophyta	61.5	68.5	XM_002503771.1	1	1548	1548
recA	<i>Ostreococcus tauri</i>	70448	Eukaryota	Viridiplantae	Chlorophyta	59.2	61.3	XM_003081668.1	1	1191	1191
recA	<i>Arabidopsis thaliana</i>	3702	Eukaryota	Viridiplantae	Streptophyta	36.7	43.6	NM_127507.2	206	1498	1293
recA	<i>Arabidopsis thaliana</i>	3702	Eukaryota	Viridiplantae	Streptophyta	36.7	46.1	NM_106556.2	40	1359	1320
recA	<i>Beta vulgaris</i> subsp. <i>vulgaris</i>	3555	Eukaryota	Viridiplantae	Streptophyta	36.5	44.5	XM_010696552.1	175	1410	1236
recA	<i>Beta vulgaris</i> subsp. <i>vulgaris</i>	3555	Eukaryota	Viridiplantae	Streptophyta	36.5	45.8	XM_010698214.1	272	1525	1254
recA	<i>Brachypodium distachyon</i>	15368	Eukaryota	Viridiplantae	Streptophyta	46.4	48.7	XM_003561970.3	125	1363	1239
recA	<i>Brachypodium distachyon</i>	15368	Eukaryota	Viridiplantae	Streptophyta	46.4	46.8	XM_003564806.3	177	1466	1290
recA	<i>Brassica napus</i>	3708	Eukaryota	Viridiplantae	Streptophyta	37.6	46.2	XM_013821439.1	97	1413	1317
recA	<i>Brassica napus</i>	3708	Eukaryota	Viridiplantae	Streptophyta	37.6	46.9	XM_013808989.1	133	1440	1308
recA	<i>Camelina sativa</i>	90675	Eukaryota	Viridiplantae	Streptophyta	37.4	46.8	XM_010431472.1	102	1415	1314
recA	<i>Camelina sativa</i>	90675	Eukaryota	Viridiplantae	Streptophyta	37.4	44.4	XM_010491192.1	91	1395	1305
recA	<i>Cicer arietinum</i>	3827	Eukaryota	Viridiplantae	Streptophyta	32.4	42.4	XM_004515991.2	1	1218	1218
recA	<i>Cicer arietinum</i>	3827	Eukaryota	Viridiplantae	Streptophyta	32.4	41.9	XM_004507889.2	125	1231	1107
recA	<i>Citrus sinensis</i>	2711	Eukaryota	Viridiplantae	Streptophyta	36.4	44.8	XM_006477935.2	84	1406	1323
recA	<i>Citrus sinensis</i>	2711	Eukaryota	Viridiplantae	Streptophyta	36.4	43.6	XM_006475411.1	305	1588	1284
recA	<i>Cucumis sativus</i>	3659	Eukaryota	Viridiplantae	Streptophyta	36.1	44.7	XM_004150973.2	44	1297	1254
recA	<i>Cucumis sativus</i>	3659	Eukaryota	Viridiplantae	Streptophyta	36.1	43.5	XM_004141009.1	230	1498	1269
recA	<i>Elaeis guineensis</i>	51953	Eukaryota	Viridiplantae	Streptophyta	41.0	46.1	XM_010906342.1	121	1362	1242
recA	<i>Elaeis guineensis</i>	51953	Eukaryota	Viridiplantae	Streptophyta	41.0	45.2	XM_010911390.1	90	1343	1254

recA	<i>Fragaria vesca</i> subsp. <i>vesca</i>	101020	Eukaryota	Viridiplantae	Streptophyta	38.9	46.4	XM_004288544.2	242	1498	1257
recA	<i>Fragaria vesca</i> subsp. <i>vesca</i>	101020	Eukaryota	Viridiplantae	Streptophyta	38.9	44.4	XM_004287523.2	176	1438	1263
recA	<i>Glycine max</i>	3847	Eukaryota	Viridiplantae	Streptophyta	35.0	42.1	XM_003546585.3	131	1375	1245
recA	<i>Glycine max</i>	3847	Eukaryota	Viridiplantae	Streptophyta	35.0	43.6	XM_014768305.1	266	1381	1116
recA	<i>Gossypium raimondii</i>	29730	Eukaryota	Viridiplantae	Streptophyta	33.7	43.9	XM_012598083.1	169	1431	1263
recA	<i>Gossypium raimondii</i>	29730	Eukaryota	Viridiplantae	Streptophyta	33.7	42.9	XM_012609242.1	249	1508	1260
recA	<i>Medicago truncatula</i>	3880	Eukaryota	Viridiplantae	Streptophyta	34.0	41.1	XM_013590005.1	137	1393	1257
recA	<i>Medicago truncatula</i>	3880	Eukaryota	Viridiplantae	Streptophyta	34.0	43.5	XM_003595157.2	176	1450	1275
recA	<i>Musa acuminata</i> subsp. <i>malaaccensis</i>	214687	Eukaryota	Viridiplantae	Streptophyta	40.7	47.0	XM_009390684.1	167	1396	1230
recA	<i>Musa acuminata</i> subsp. <i>malaaccensis</i>	214687	Eukaryota	Viridiplantae	Streptophyta	40.7	45.7	XM_009383564.1	133	1389	1257
recA	<i>Oryza sativa</i> Japonica Group	39947	Eukaryota	Viridiplantae	Streptophyta	43.5	51.3	NM_001057273.1	44	1300	1257
recA	<i>Oryza sativa</i> Japonica Group	39947	Eukaryota	Viridiplantae	Streptophyta	43.5	47.1	NM_001051644.1	104	1393	1290
recA	<i>Phaseolus vulgaris</i>	3885	Eukaryota	Viridiplantae	Streptophyta	36.9	42.1	XM_007133257.1	125	1372	1248
recA	<i>Phaseolus vulgaris</i>	3885	Eukaryota	Viridiplantae	Streptophyta	36.9	41.5	XM_007154579.1	138	1220	1083
recA	<i>Populus trichocarpa</i>	3694	Eukaryota	Viridiplantae	Streptophyta	34.1	42.4	XM_006370406.1	130	1413	1284
recA	<i>Populus trichocarpa</i>	3694	Eukaryota	Viridiplantae	Streptophyta	34.1	44.1	XM_002308248.2	262	1518	1257
recA	<i>Prunus mume</i>	102107	Eukaryota	Viridiplantae	Streptophyta	38.3	44.0	XM_008229689.1	130	1419	1290
recA	<i>Prunus mume</i>	102107	Eukaryota	Viridiplantae	Streptophyta	38.3	44.1	XM_008244142.1	113	1360	1248
recA	<i>Sesamum indicum</i>	4182	Eukaryota	Viridiplantae	Streptophyta	35.2	44.5	XM_011097372.1	112	1377	1266
recA	<i>Sesamum indicum</i>	4182	Eukaryota	Viridiplantae	Streptophyta	35.2	45.1	XM_011079694.1	306	1577	1272
recA	<i>Setaria italica</i>	4555	Eukaryota	Viridiplantae	Streptophyta	46.2	49.3	XM_004982290.1	65	1318	1254
recA	<i>Setaria italica</i>	4555	Eukaryota	Viridiplantae	Streptophyta	46.2	46.4	XM_004970885.3	218	1501	1284
recA	<i>Solanum lycopersicum</i>	4081	Eukaryota	Viridiplantae	Streptophyta	35.7	44.0	XM_004243946.2	158	1417	1260
recA	<i>Solanum lycopersicum</i>	4081	Eukaryota	Viridiplantae	Streptophyta	35.7	42.4	XM_010323031.1	57	1331	1275
recA	<i>Sorghum bicolor</i>	4558	Eukaryota	Viridiplantae	Streptophyta	44.3	49.4	XM_002466721.1	60	1307	1248
recA	<i>Sorghum bicolor</i>	4558	Eukaryota	Viridiplantae	Streptophyta	44.3	46.7	XM_002458886.1	172	1455	1284
recA	<i>Theobroma cacao</i>	3641	Eukaryota	Viridiplantae	Streptophyta	35.2	42.8	XM_007022241.1	556	2013	1458
recA	<i>Theobroma cacao</i>	3641	Eukaryota	Viridiplantae	Streptophyta	35.2	42.1	XM_007012745.1	417	1679	1263
recA	<i>Vigna radiata</i> var. <i>radiata</i>	3916	Eukaryota	Viridiplantae	Streptophyta	34.1	42.2	XM_014637318.1	103	1353	1251
recA	<i>Vigna radiata</i> var. <i>radiata</i>	3916	Eukaryota	Viridiplantae	Streptophyta	34.1	41.8	XM_014651708.1	176	1432	1257
recA	<i>Vitis vinifera</i>	29760	Eukaryota	Viridiplantae	Streptophyta	35.0	45.8	XM_002271691.3	196	1443	1248
recA	<i>Vitis vinifera</i>	29760	Eukaryota	Viridiplantae	Streptophyta	35.0	44.1	XM_002279428.3	196	1461	1266
recA	<i>Zea mays</i>	381124	Eukaryota	Viridiplantae	Streptophyta	46.8	48.9	XM_008678176.1	154	1401	1248
recA	<i>Zea mays</i>	381124	Eukaryota	Viridiplantae	Streptophyta	46.8	44.5	XM_008675034.1	1	1365	1365
recA	<i>Microcystis aeruginosa</i> phage Ma-LMM01	340435	Viruses	dsDNA viruses, no RNA stage	Caudovirales	46.0	47.3	NC_008562.1	7777	8841	1065
uvvX	<i>Acinetobacter</i> phage 133	279006	Viruses	dsDNA viruses, no RNA stage	Caudovirales	39.7	40.8	NC_015250.1	32171	33334	1164
uvvX	<i>Aeromonas</i> phage CC2	1204516	Viruses	dsDNA viruses, no RNA stage	Caudovirales	38.8	40.3	NC_019538.1	99753	100862	1110
uvvX	<i>Dickeya</i> phage vB_DsoM_LIMEstone1	1091052	Viruses	dsDNA viruses, no RNA stage	Caudovirales	49.3	51.7	NC_019925.1	30946	32031	1086
uvvX	<i>Enterobacteria</i> phage RB14	197310	Viruses	dsDNA viruses, no RNA stage	Caudovirales	35.3	35.9	NC_012638.1	20614	21795	1182
uvvX	<i>Enterobacteria</i> phage vB_EcoM-VR7	700939	Viruses	dsDNA viruses, no RNA stage	Caudovirales	40.3	41.2	NC_014792.1	22953	24131	1179
uvvX	<i>Erwinia</i> phage PhiEaH1	1401669	Viruses	dsDNA viruses, no RNA stage	Caudovirales	52.3	52.2	NC_023610.1	142257	143684	1428
uvvX	<i>Escherichia</i> phage Lw1	1307804	Viruses	dsDNA viruses, no RNA stage	Caudovirales	43.5	44.5	NC_021344.2	21151	22296	1146
uvvX	<i>Klebsiella</i> phage KP27	1129147	Viruses	dsDNA viruses, no RNA stage	Caudovirales	41.8	43.2	NC_020080.1	20378	21535	1158
uvvX	<i>Prochlorococcus</i> phage P-SSM4	268747	Viruses	dsDNA viruses, no RNA stage	Caudovirales	36.7	37.6	NC_006884.2	146410	147420	1011
uvvX	<i>Serratia</i> phage PS2	1481112	Viruses	dsDNA viruses, no RNA stage	Caudovirales	41.7	43.2	NC_024121.1	20481	21674	1194
uvvX	<i>Shigella</i> phage SP18	645664	Viruses	dsDNA viruses, no RNA stage	Caudovirales	40.4	41.3	NC_014595.1	21866	23044	1179
uvvX	<i>Sinorhizobium</i> phage phiIM12	1357423	Viruses	dsDNA viruses, no RNA stage	Caudovirales	49.0	50.0	NC_027204.1	8976	10043	1068
uvvX	<i>Synechococcus</i> phage S-CRM01	1026955	Viruses	dsDNA viruses, no RNA stage	Caudovirales	39.7	39.3	NC_015569.1	109534	110547	1014
uvvX	<i>Synechococcus</i> phage S-RSM4	555387	Viruses	dsDNA viruses, no RNA stage	Caudovirales	41.1	44.8	NC_013085.1	46738	47754	1017
uvvX	<i>Vibrio</i> phage nt-1	115992	Viruses	dsDNA viruses, no RNA stage	Caudovirales	41.3	41.2	NC_021529.2	168873	169973	1101
uvvX	<i>Yersinia</i> phage PST	1351740	Viruses	dsDNA viruses, no RNA stage	Caudovirales	35.3	36.1	NC_027404.1	20520	21701	1182

# 相同組換え酵素遺伝子の系統樹 詳細

第3章に掲載した相同組換え酵素遺伝子の系統樹の元となった、各枝を束ねていない分子系統樹を掲載する。





**recA**

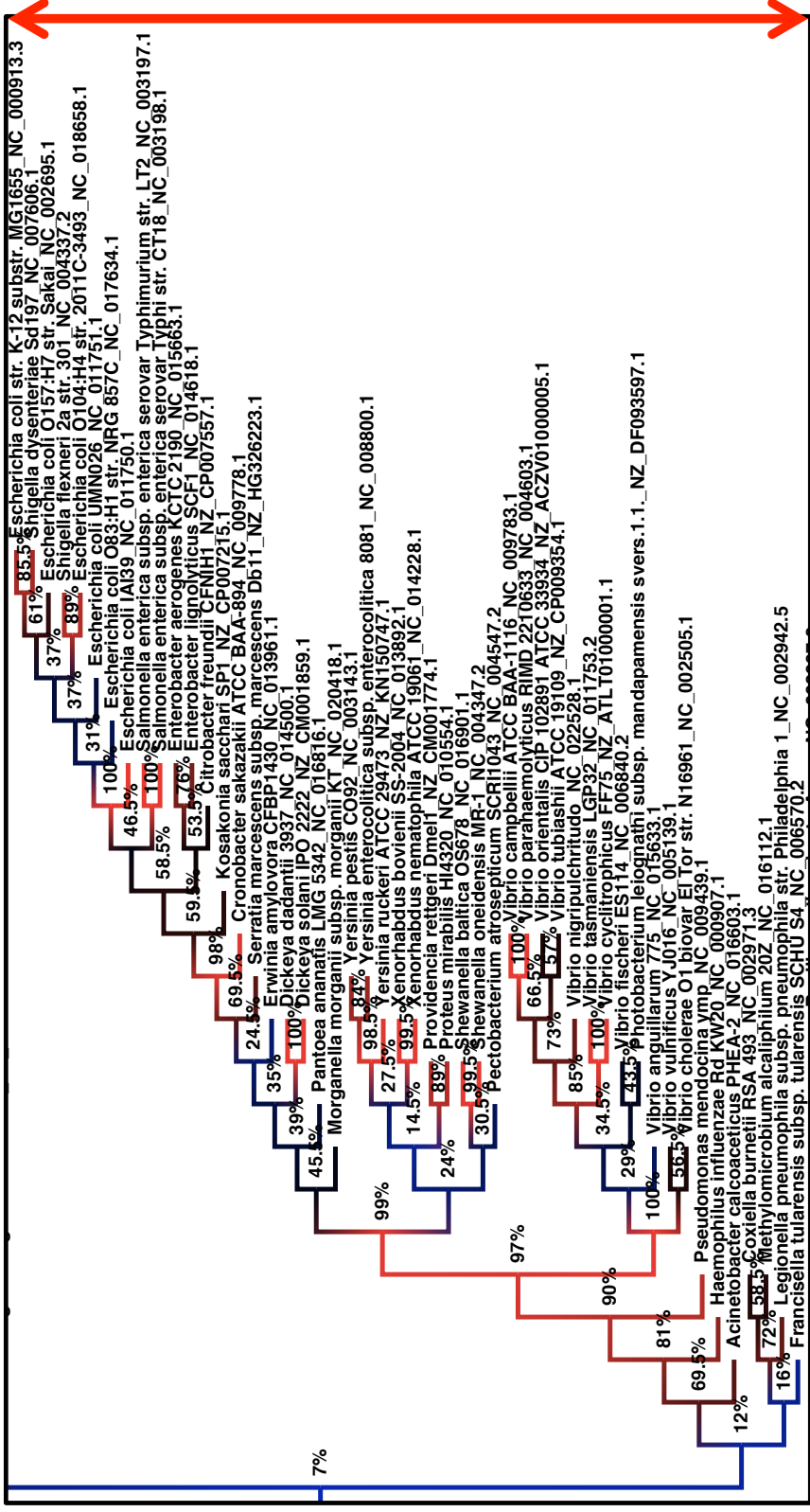
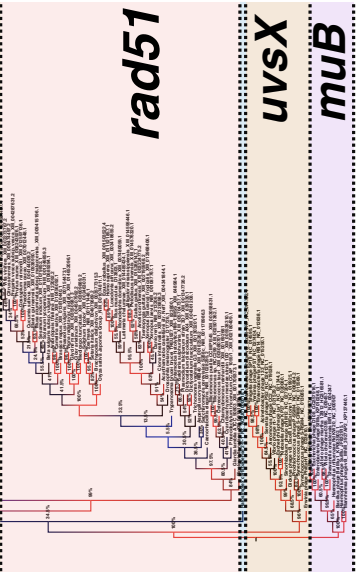
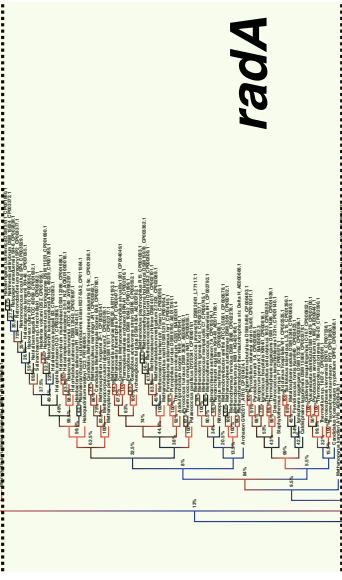
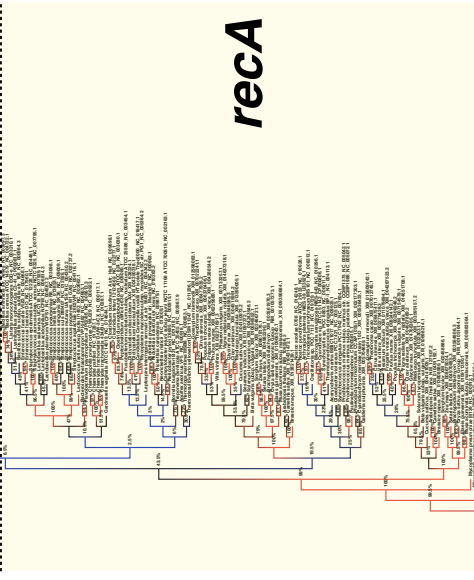
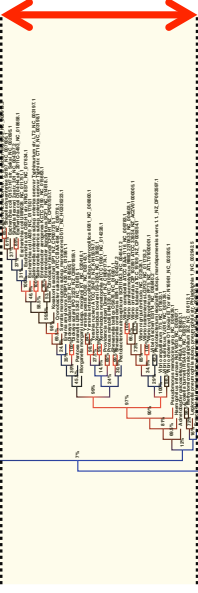
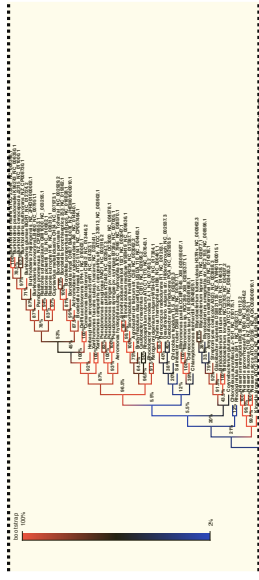
**radA**

**rad51**

**kaIC**

**uvrX**

**muB**



recA

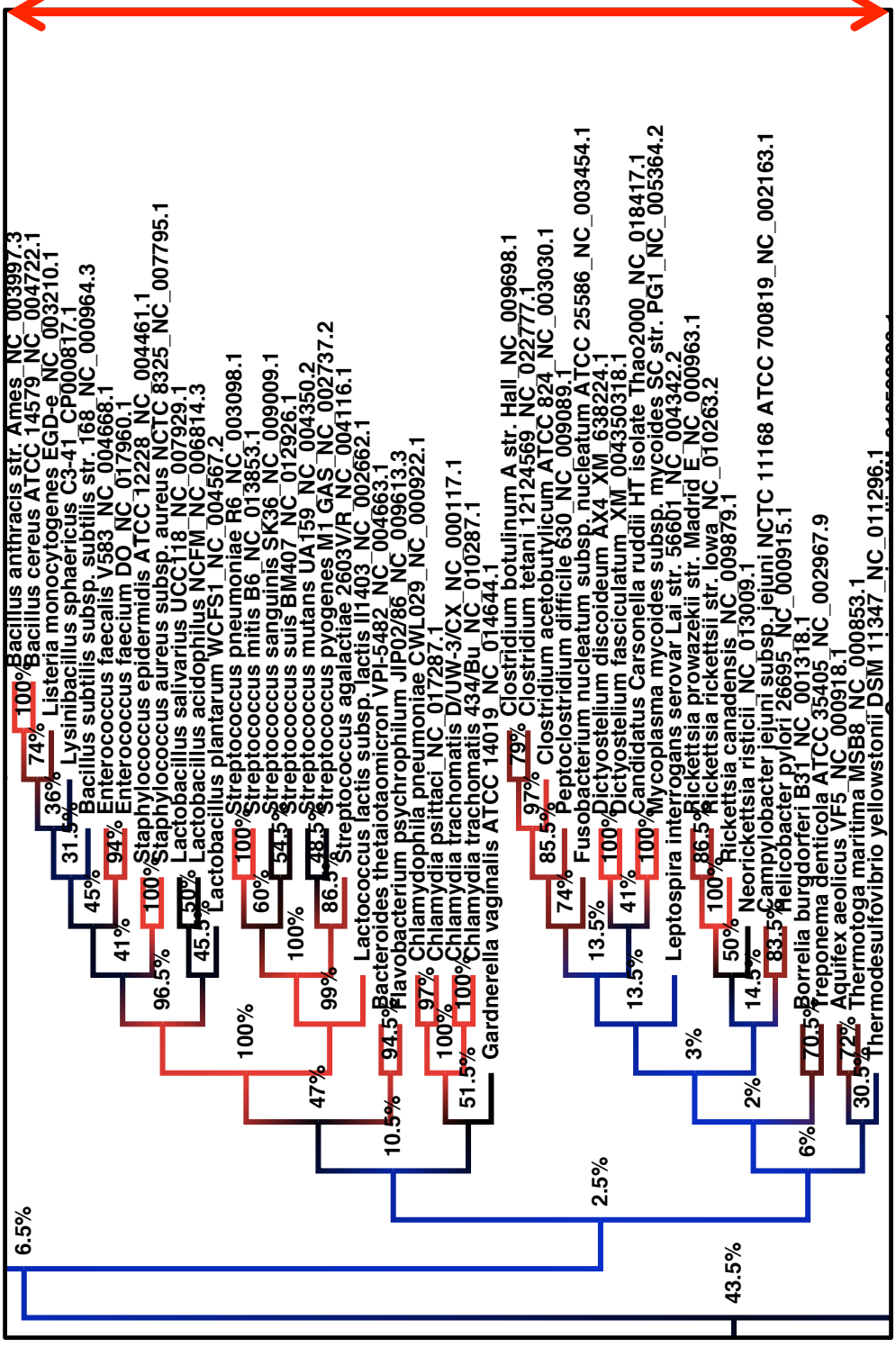
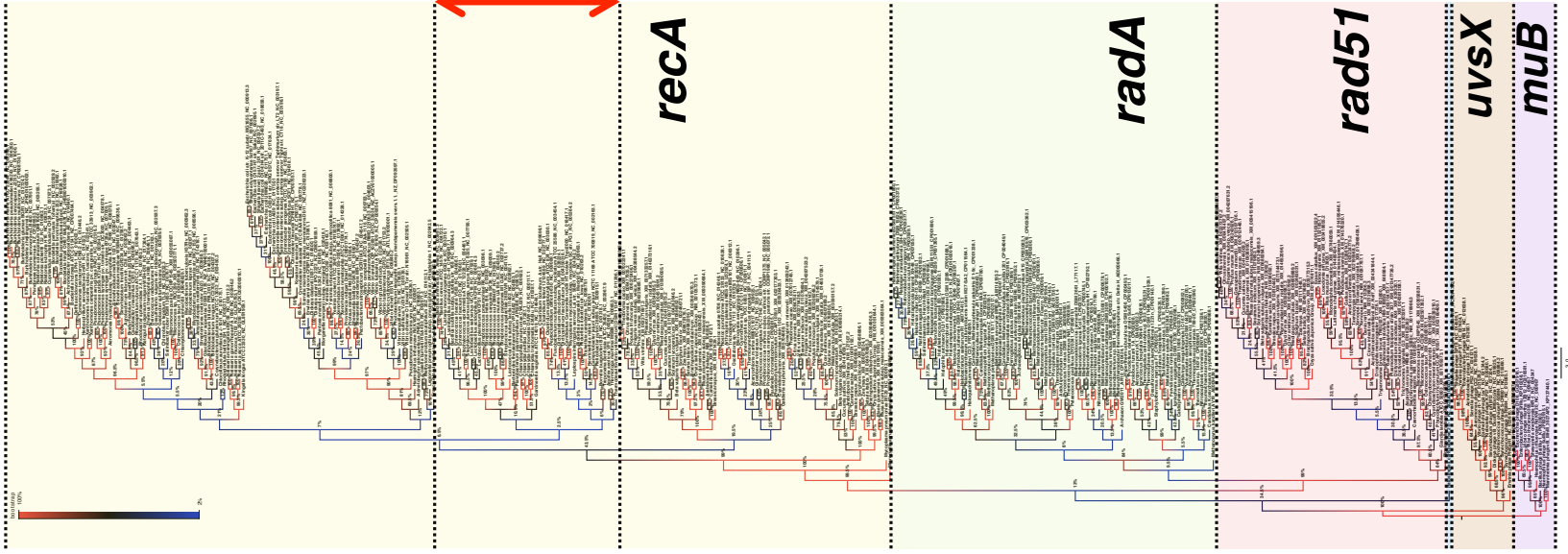
radA

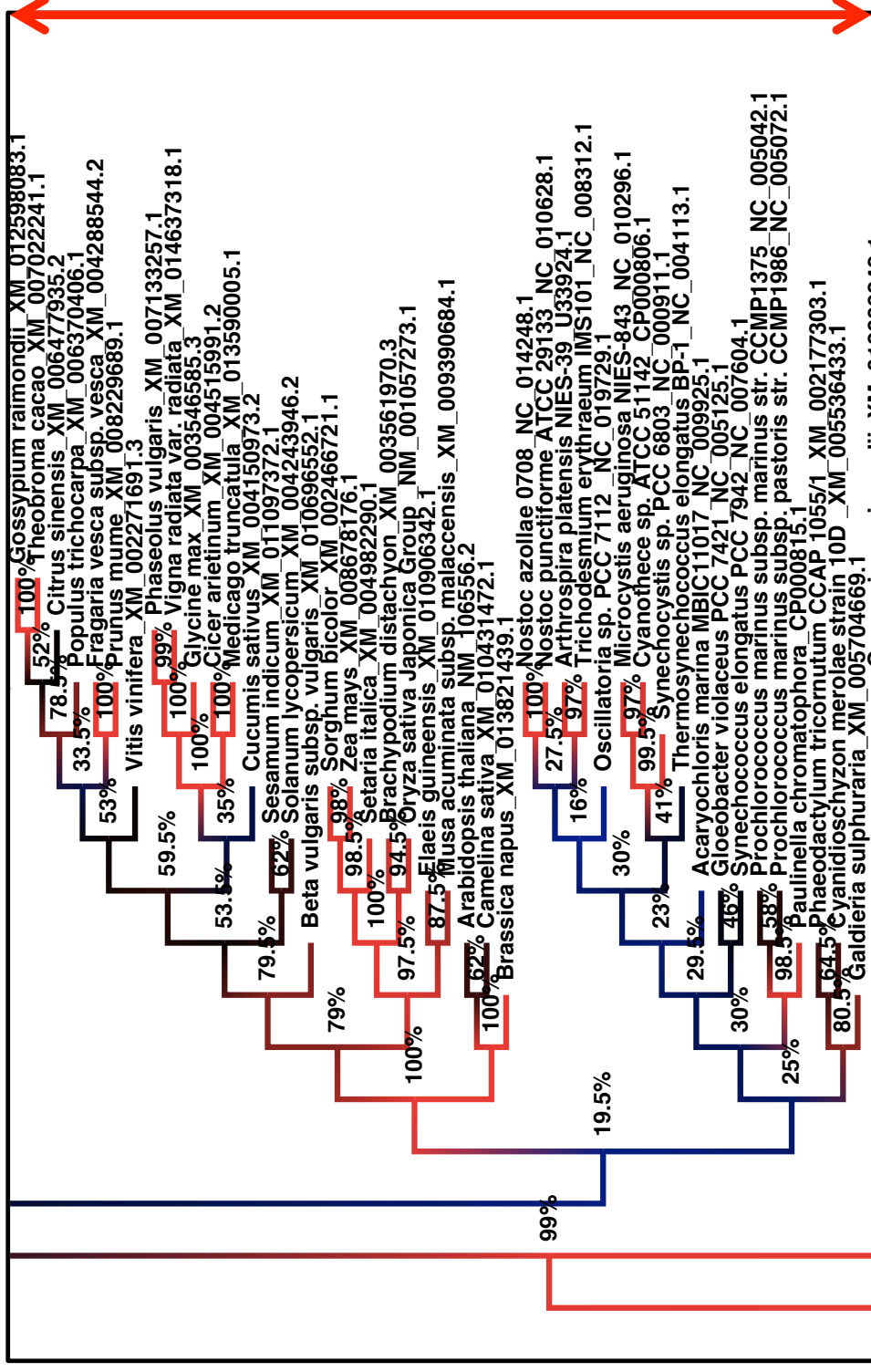
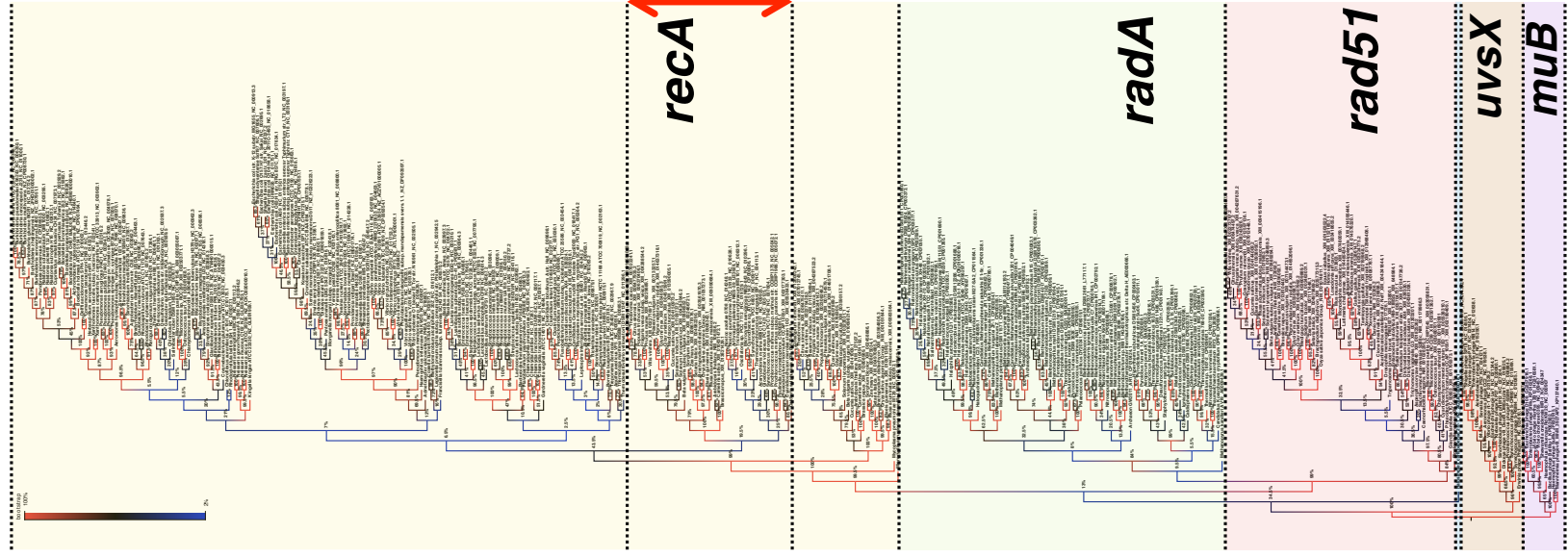
rad51

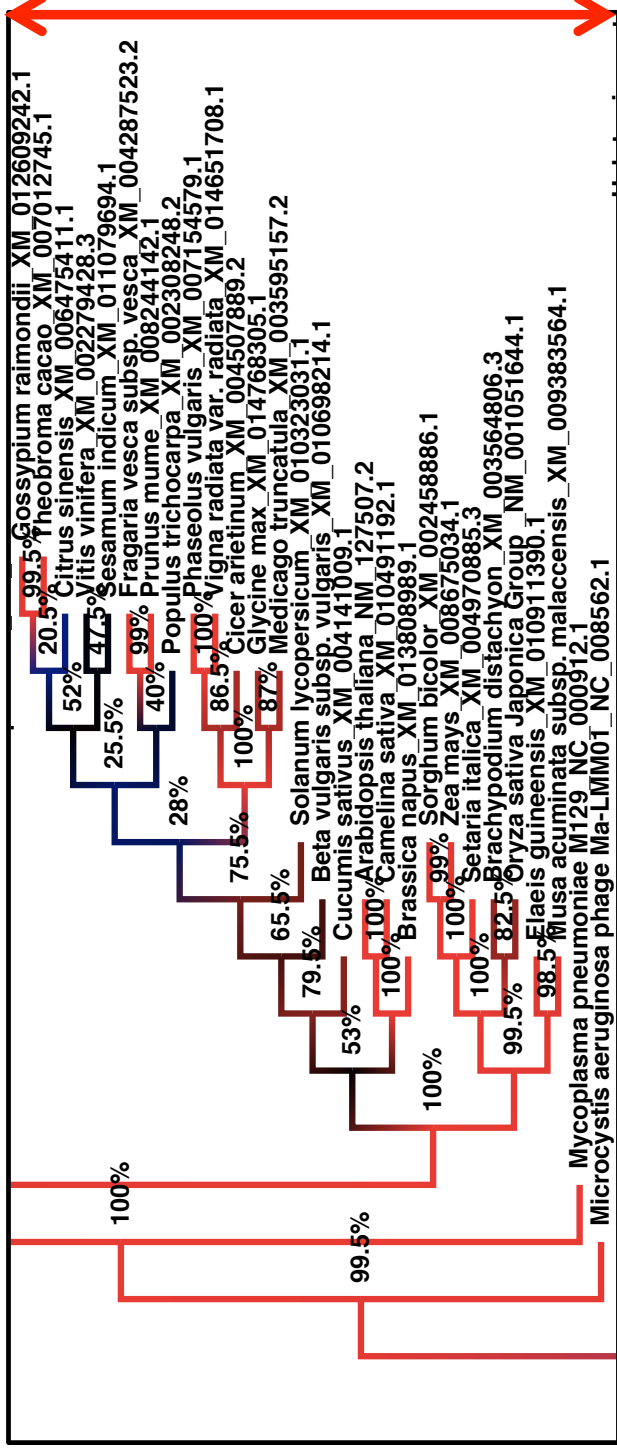
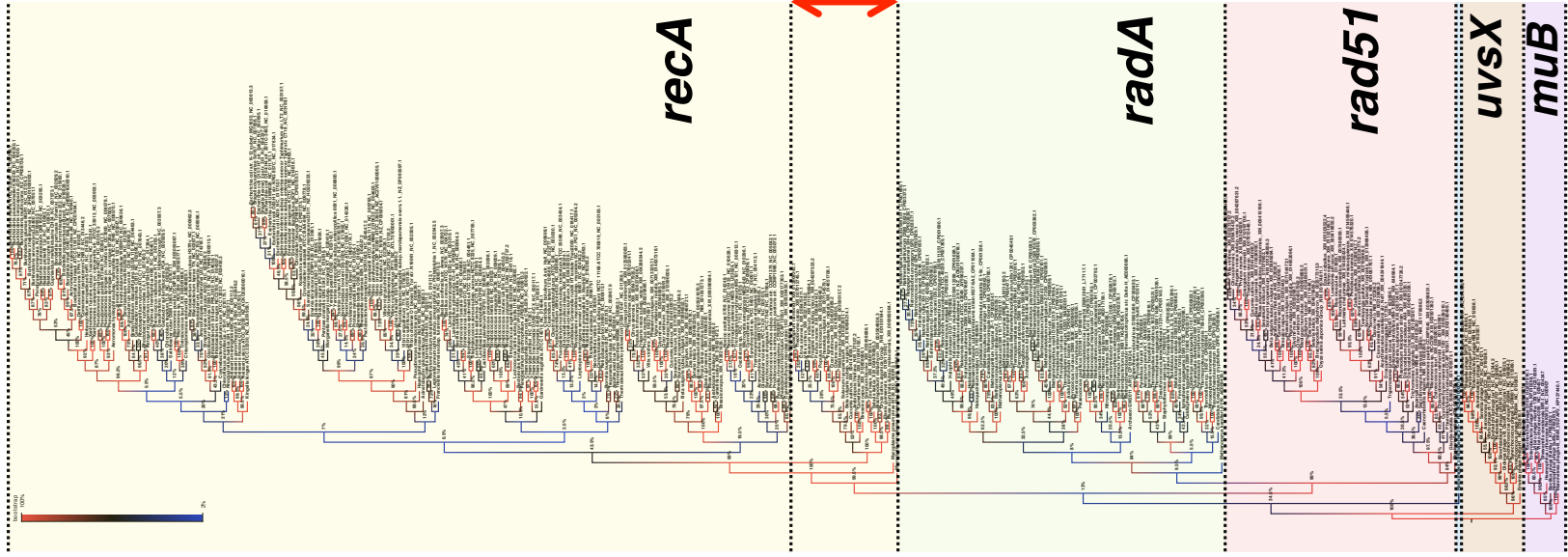
kaic

uvsX

muB



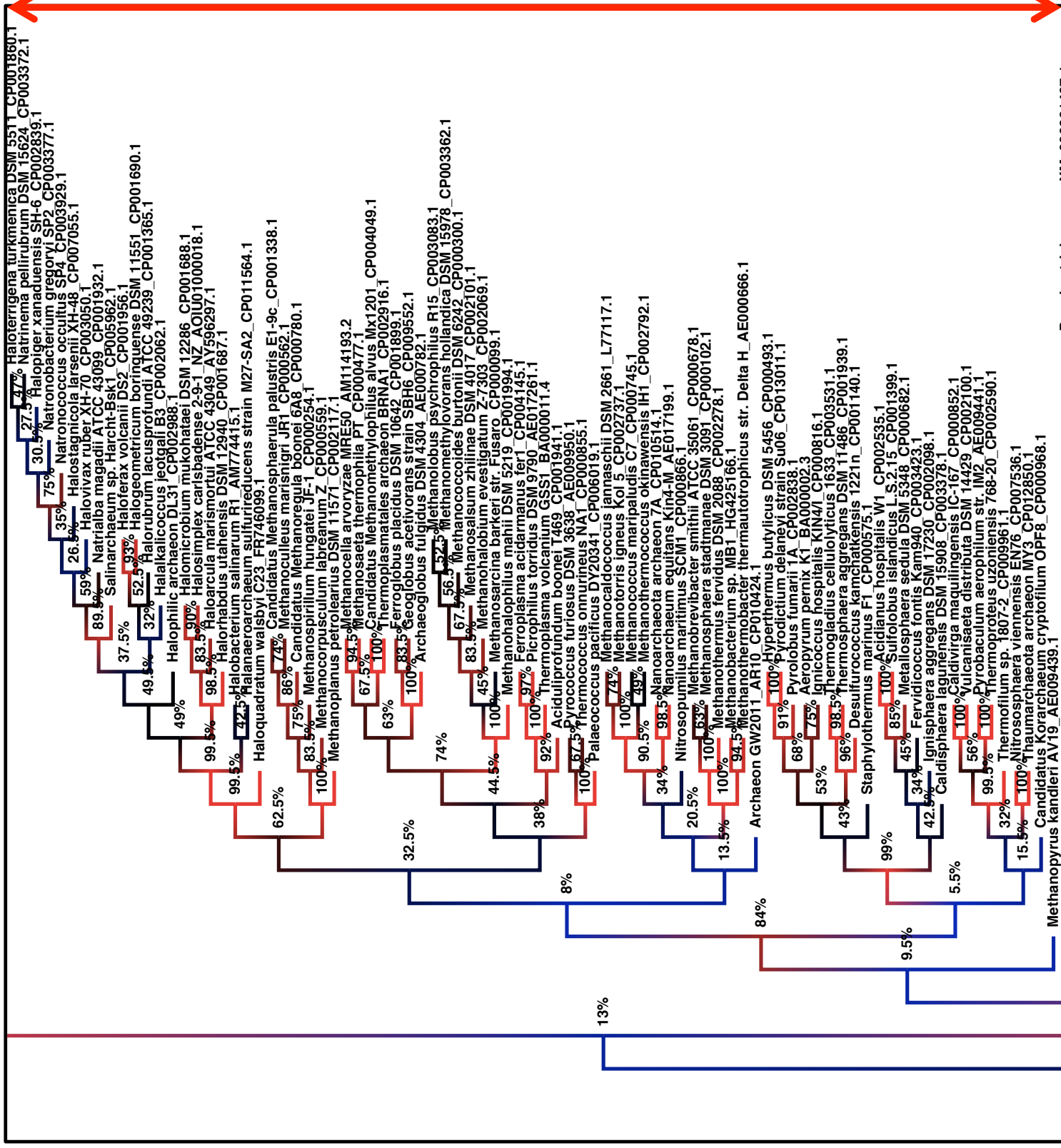
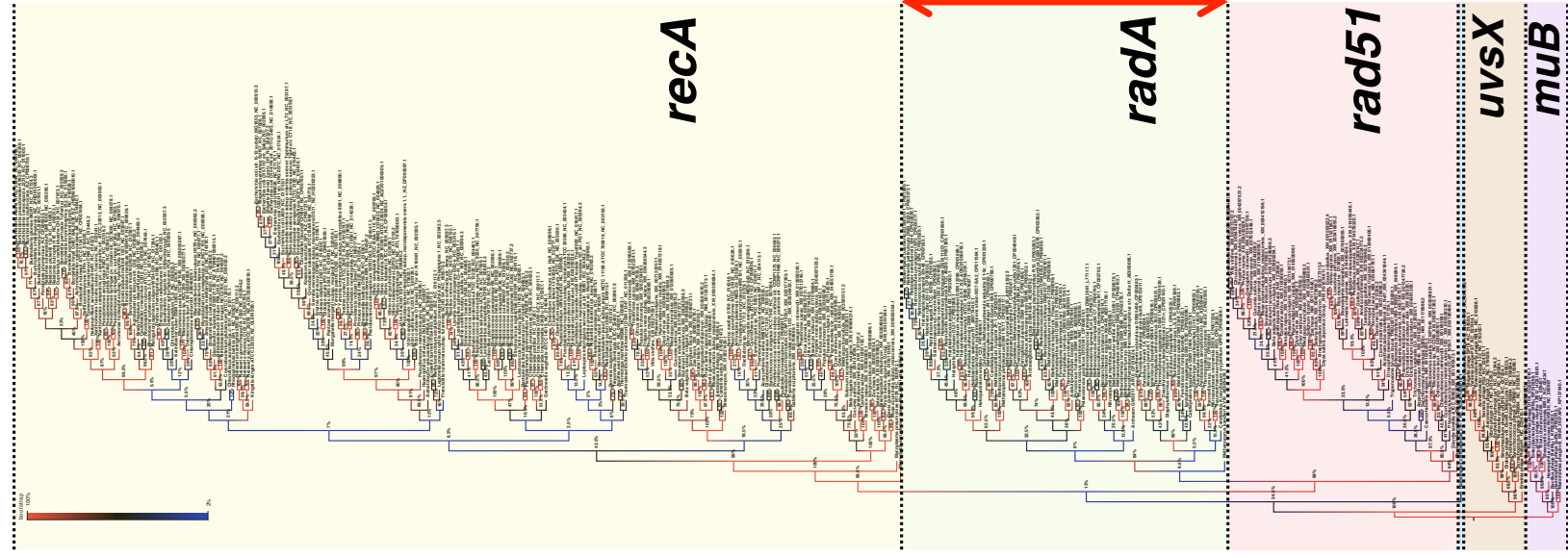




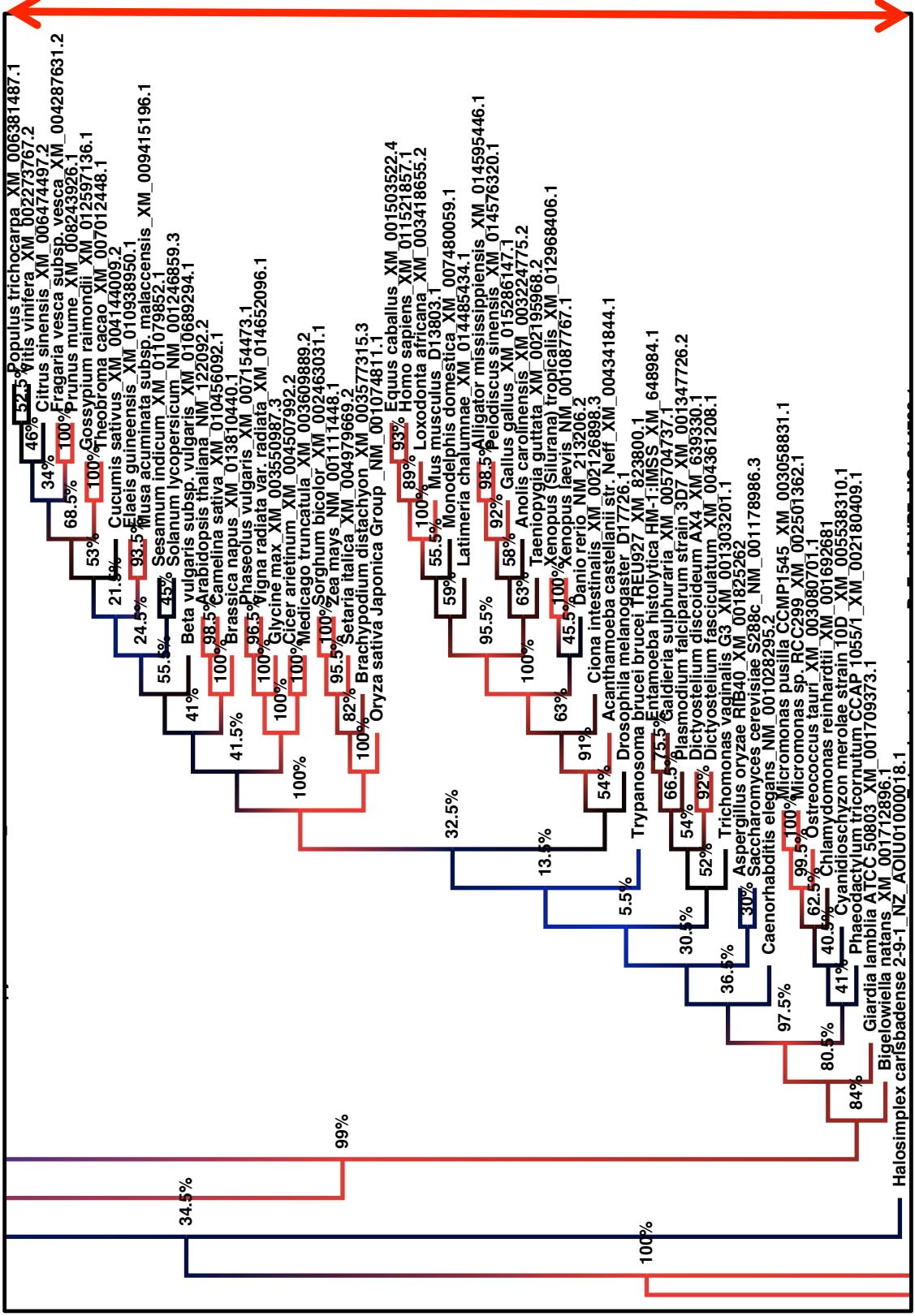
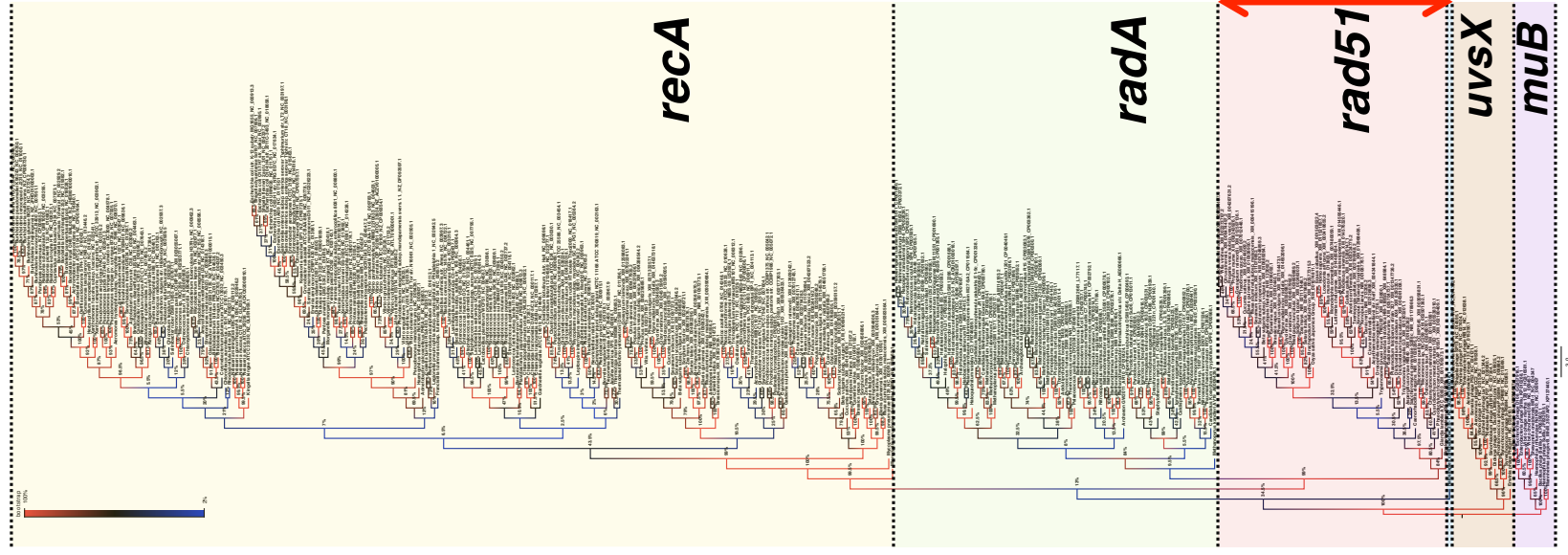
*kaiC*

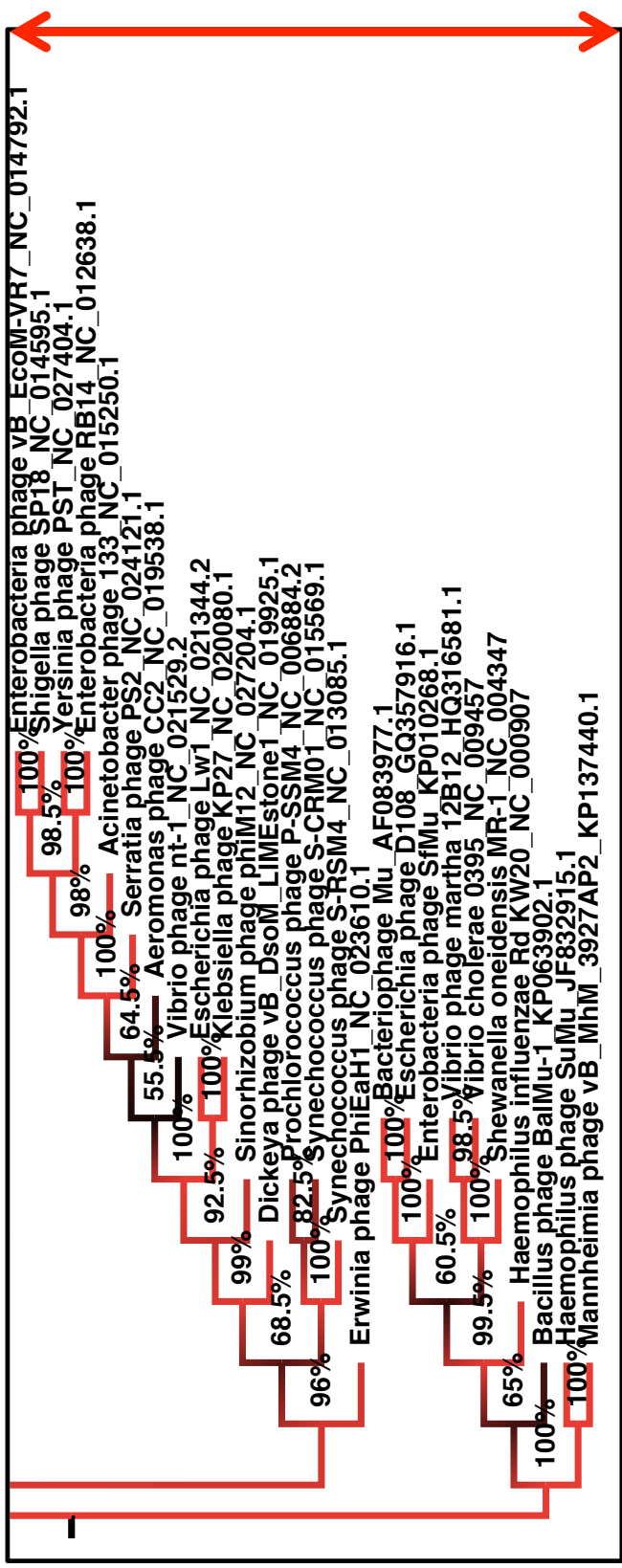
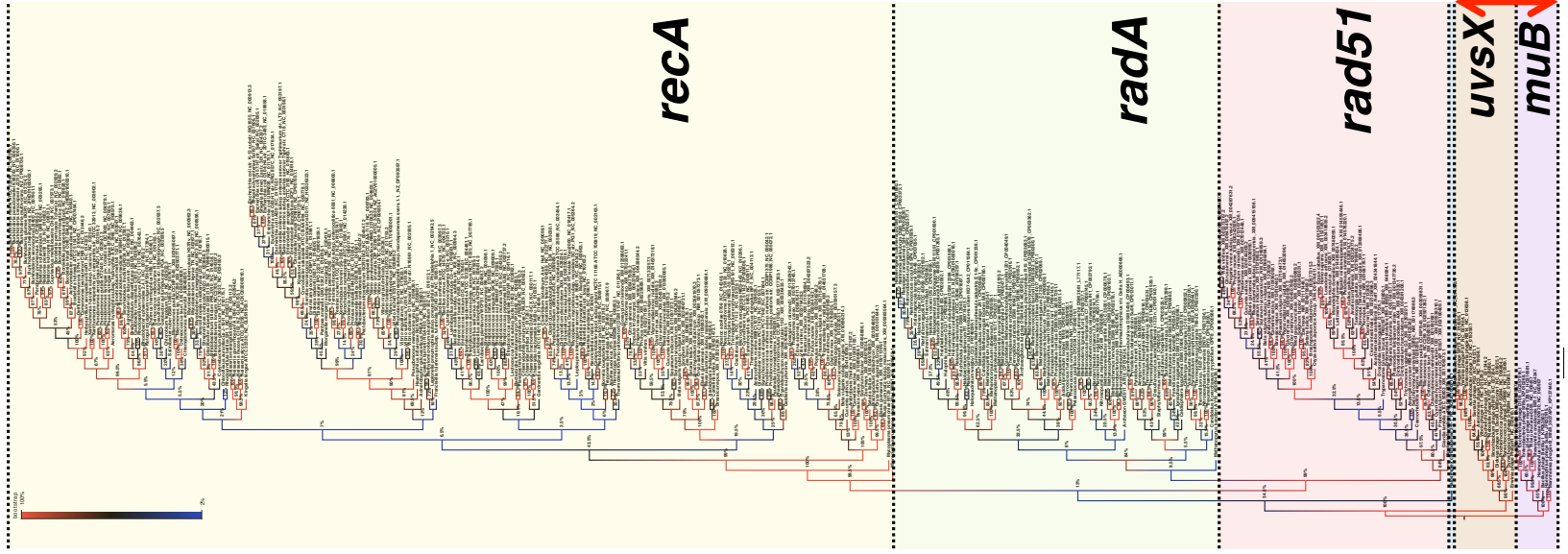
*uvsX*

*muB*



kaiC







# 分子系統解析と塩基配列のGC含量の 関係性解析

第3章に関連し、一般に塩基配列のGC含量と分子系統樹の間に相関性が生じるか、シミュレーションにより確認した結果を掲載する。

第3章, 第1節において, *recA*と*radA*の分子系統樹と解析に用いた配列のGC含量に相関性が見られることを示した(図9)。この結果が, 分子系統解析に依存して生じた結果でないことを示すために, シミュレーションを行った。ある親配列を元に塩基配列がランダムに変化した配列群を作製し, 分子系統解析を行った場合, 系統関係と配列のGC含量に相関性が見出されるか, 以下の通り確認した。

#### <方法>

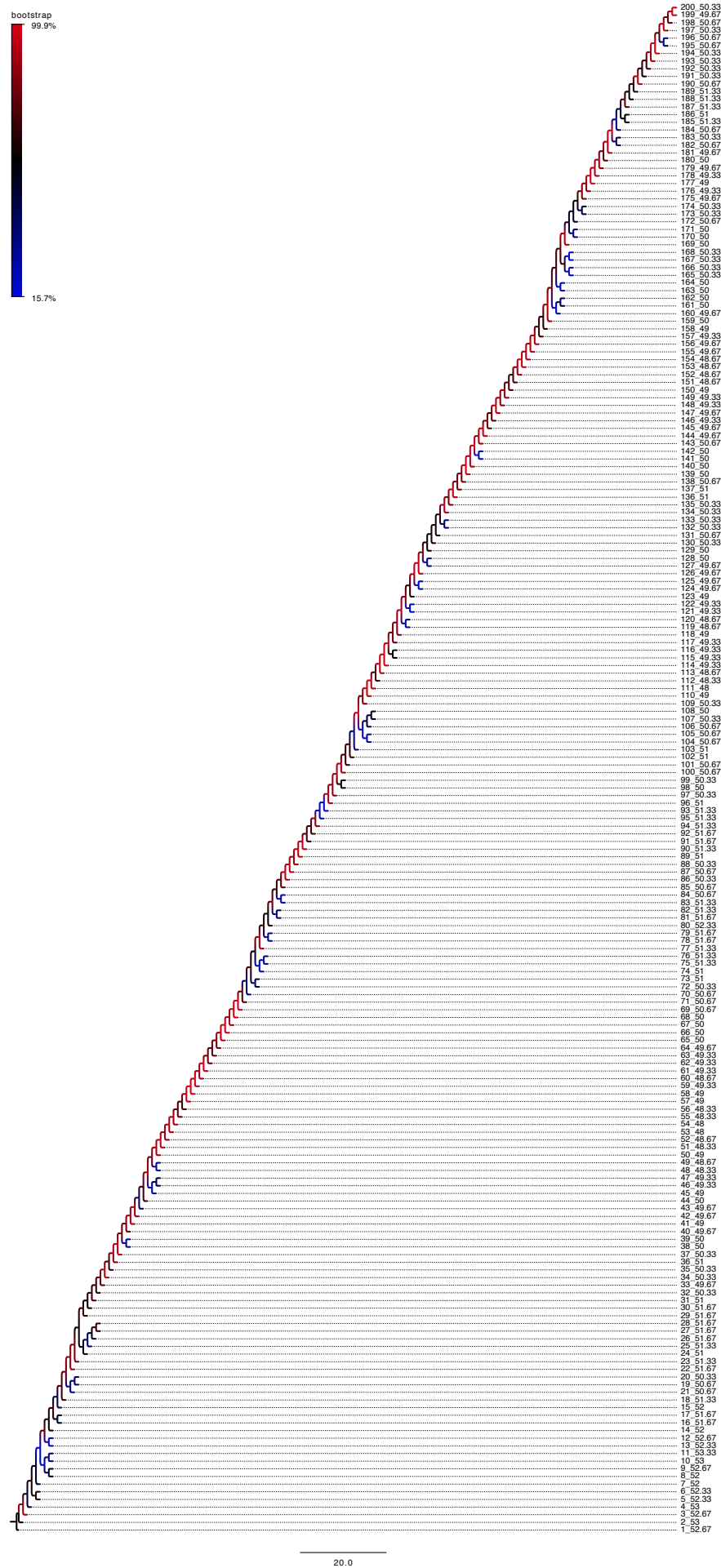
1. 300 bpからなる塩基配列を生成した。これを親配列とした(1世代目)。ただし, 配列はランダムである。
2. 生成した配列を参照し, 次世代の配列を生成した。ただし, 各座位が1%の確率で変異が起き, かつ75%の確率で元の塩基とは異なる塩基に置換されるものとした。従って, 各座位において塩基置換を伴う変異が生じる確率は0.75%である。
3. 2. のプロセスを200回繰り返し, 塩基配列を200本生成した。
4. 得られた200本の配列のGC含量を各々計算した。
5. 200本の配列を整列させ, 近隣接合法 (Neighbor-Joining法) を用いて系統樹を作製した。系統樹は1000回の試行によるbootstrap確率を計算し, この値に基づくものを使用した (Bootstrap-consensused phylogenetic tree)。なお, 分子系統樹の作製にはMEGA7 (version 7.0.14) を使用した[1]。また, 系統樹の描画にはFigTree (version 1.4.2) を使用した (<http://tree.bio.ed.ac.uk/software/figtree/>)。
6. 分子系統樹の枝に番号を振り, これらの値に対して各配列のGC含量をプロットした。

## <結果, 考察>

作製した塩基配列を用いて作製した系統樹を示す（付録 4, 図1）. 塩基配列は一つ前の世代を参照して変異するため, 得られた系統樹はほぼ単調に子ノードが形成された. 配列の世代数とノードの関係にも, ほぼ矛盾は見られない. つまり, 系統解析により得られる情報は, 世代に伴う配列の変化を反映し得るものであることが示された. 一方で, この系統樹の枝分かれに対して各配列のGC含量をプロットしたが, 系統樹の枝番号の増加に対し, 系統関係に対するGC含量の変化は単調に増加する様子は見られず, 振動していることが分かった（付録 4, 図2）. このことから, 系統樹の枝分かれは, 配列のGC含量に影響されないことが分かった. 第3章, 第1節, 図9と同様にして, 図8の系統樹全体についてGC含量をプロットした図を示す（付録 4, 図3）. 図から明らかのように, バクテリアの*recA*や一部の古細菌の*radA*クレードではGC含量が線型に増加する様子が確認できるが, 系統樹全体で単調に増加はしていないことが見て取れる. 以上から, 分子系統樹の作製に際してGC含量の増加が解析の性質上生じたものではないことが示された. では, どのような場合にGC含量と系統関係に正または負の相関性が見られるのだろうか. シミュレーションにより配列を生成する際, 塩基配列の置換確率に偏りを持たせれば, 世代を重ねるごとに生成される配列は重み付けされた塩基に集約されていくことは自明である. 生成された塩基配列を用いて系統樹を作製しても, その系統関係は塩基配列の世代をほぼ反映することは既に示した通りである（付録 4, 図1）. つまり, 塩基配列の生成段階で塩基置換の確率に偏りがある場合に, 系統関係とGC含量の推移に相関性が生じるということである. これらの解析を考慮すれば, 第3章, 第1節の図9に示した*recA*及び*radA*におけるGC含量の線型な増加は, 相同組換え遺伝子の系統分岐に際する塩基置換の偏りを示唆していると言える.

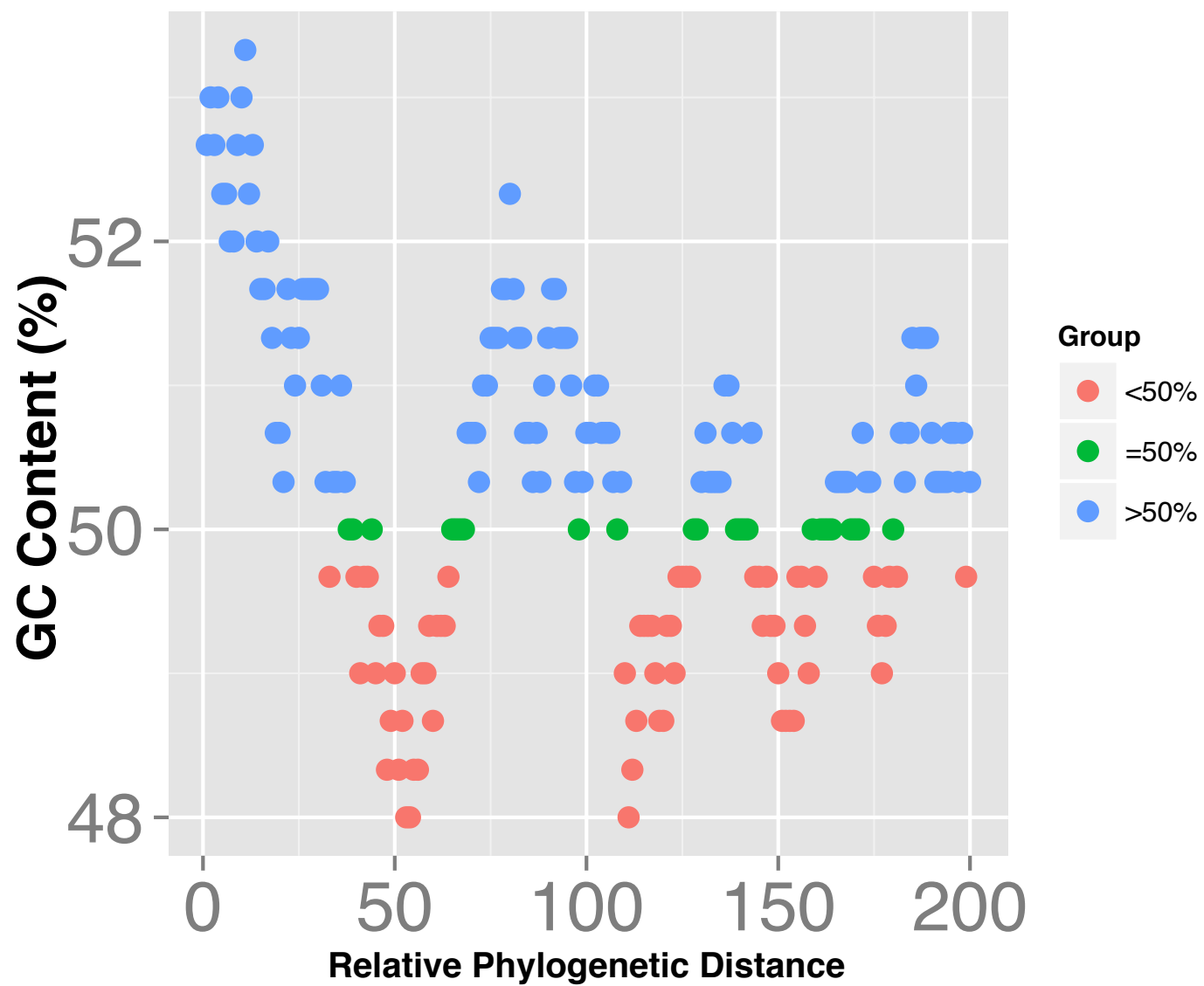
## 参考文献

1. Kumar, S., Stecher, G., and Tamura, K. MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets, *Mol. Biol. Evol.* 2016;33, 1870-1874.



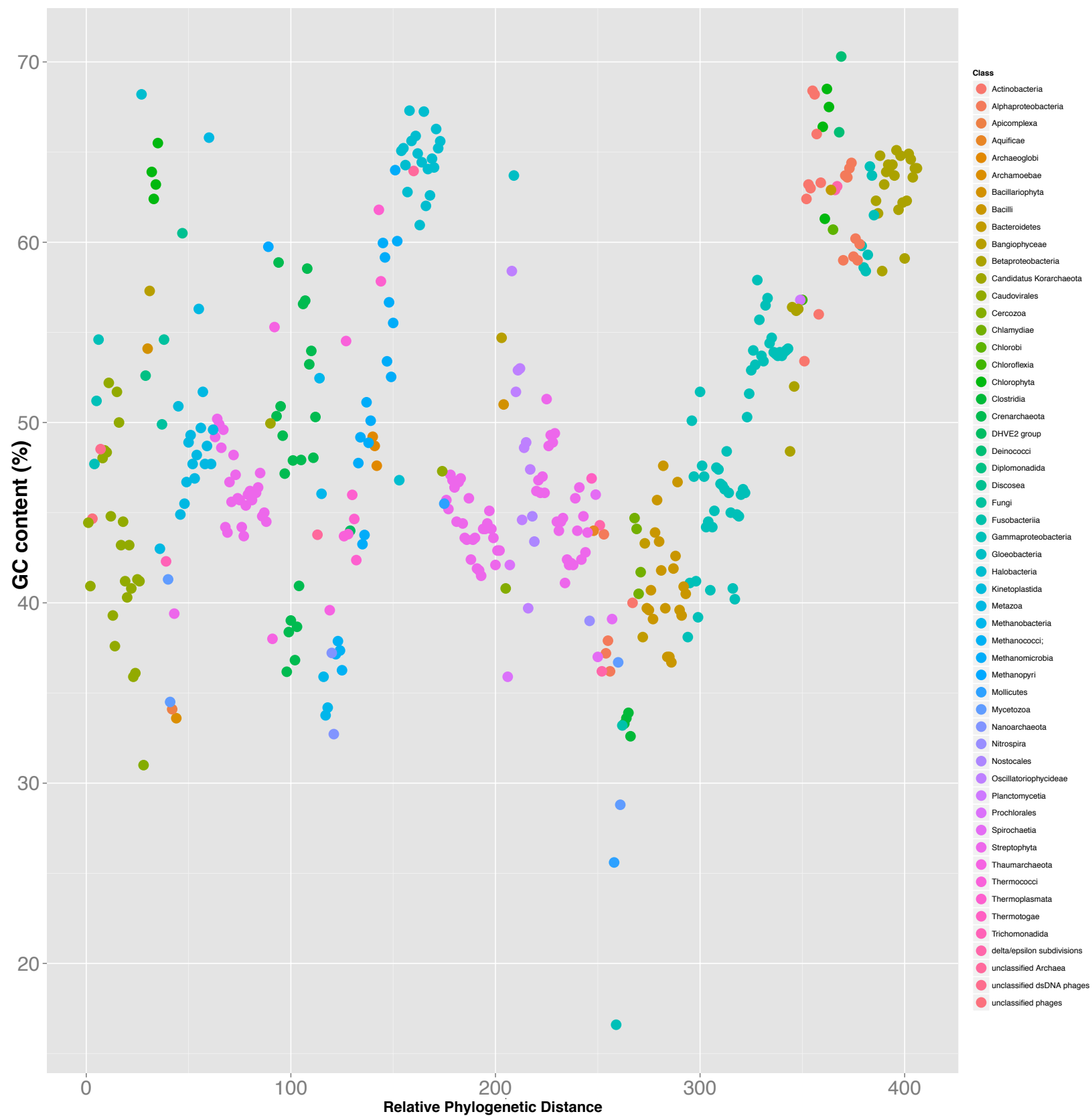
付録 4 図1. 生成した塩基配列を用いた分子系統樹.

生成した配列を用いて作製した分子系統樹を示す。ラベルの頭の値は塩基配列の世代番号, アンダーバー ( \_ ) を挟んで右側の値が塩基配列のGC含量である(i.e. 185\_51.33: 185世代目, GC含量51.33%). 左上のスケールバーはbootstrap確率を示し, 系統樹の枝の色と対応している。



付録 4 図2. 塩基配列がランダムに変異する場合における系統関係とGC含量の相関性.

付録 4 図1の系統樹の枝番号 (Relative Phylogenetic Distance) に対して塩基配列のGC含量 (GC content(%)) をプロットした図を示す. ドットの色は塩基配列のGC含量と対応しており, 赤は50%未満 (<50%), 緑は50% (=50%), 青は50%より大 (>50%) とした.



付録 4 図2. 図8で示した相同組換え遺伝子の分子系統樹に基づく、全系統関係とGC含量の相関性.

第3章, 第1節の図8で示した相同組換え遺伝子の分子系統樹に基づく、全系統関係とGC含量の相関性を示す。縦軸は解析した相同組換え遺伝子のGC含量塩基配列のGC含量 (GC content(%)) , 横軸は系統樹の枝番号 (Relative Phylogenetic Distance) を意味する。ドットの色は各OTU (Operational Taxisonomic Unit) の門 (Phylum) または綱 (Class) と対応している。

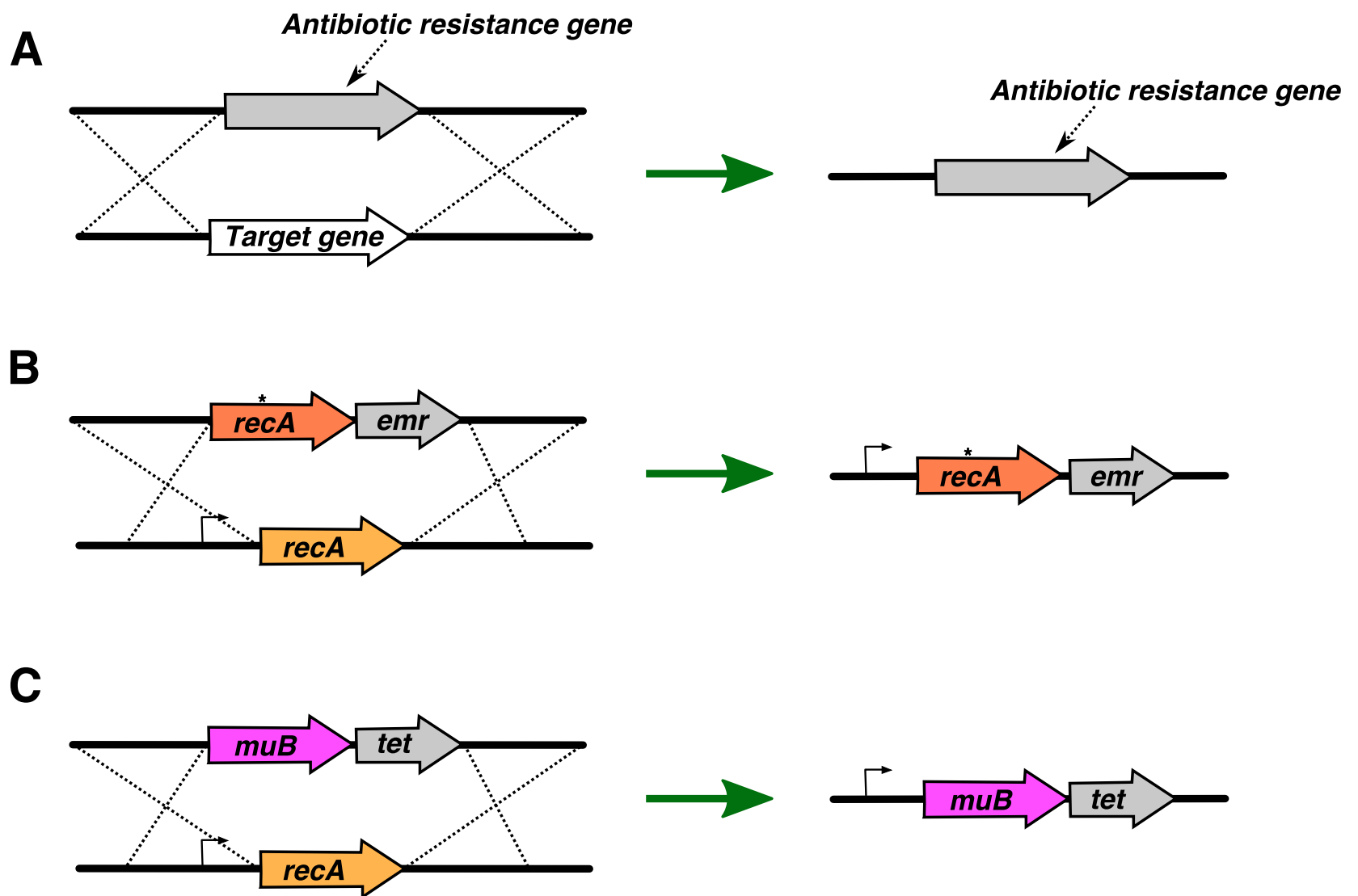
# 株の作製に用いたプライマーのリスト 及び作製した株の構造

第2章及び第3章で実験に使用したプライマーの配列名及び配列リスト，作製した株の構造を掲載する。

	Information	Primer name	Sequence (5'>3')
<b>Plasmid construction</b>			
T2	<i>glyQ/StRNA</i> attenuator	A1-F A1-R	gggcatgcGAACTCTCGTCCCTAIGTTTG TTCCACACGTTCAAGCACCTCCACTTTTC
	<i>xkd</i> operon terminator	A2-F A2-R	GTGCTTGAACGTGTGGAATTTGAAATCG gggctagcAAATAAAAAACGGACACCAATC
mini-IS	IS256 <i>BsuI</i> IRL SD <i>cat</i> foward	B1-F	GTGTAAGTAAGGAATGCGTACGTAAaaggaggtgtaaacatg
	IS256 <i>BsuI</i> IRR <i>cat</i> reverse	B1-R	GAAAAGTGTGTAAGTTATTCTAGATAATTCCTTCTCAACTAACGGG
	NheI-IS256 <i>BsuI</i> IRL foward	B2-F	gggctagcGGGAGTGTAAATTAAGTGTGTAAGTAAGGAATGCG
	Sall-IS256 <i>BsuI</i> IRR reverse	B2-R	gcgtcgacGGGCTTGTCAAGAAAAGTGTGTAAGTTATTCTAG
<i>tnp</i>	HindIII-SD-IS256 <i>BsuI tnp</i> foward	C-F	ggaagcttTAAAAGGAGGTtgtaacATGGGGAAATTGAAAAGAG
	Sall-IS256 <i>BsuI tnp</i> reverse	C-R	gggtcgacTTATTCTAGATAATTCCTCACACG
<b>Mutant construction</b>			
<i>recO::erm</i>		F1-F F1-R F2-F F2-R F3-F F3-R	CTAAGCATAAACTTGGGGACTTTATGATG agtcagcacgtaTTCCGCACCTTCTCAAGCAG gaaggtcggaaTAACGTGCTGACTTGCACCATATC gatggaccaaatTTACTATTCCTCCCGTTAAATAATAGATA gaggaaatagtaaCATTTGGTCCATCTTTTATATGATCATTATTATAAAATATG GCTGGATGAACAAGTCCGTTGTC
<i>recU::erm</i>		F1-F F1-R F2-F F2-R F3-F F3-R	GTAAATCCGACGGTGTCTCCGTC agtcagcacgtaTTTCTCATCCTCCTTATTAATAATCATGAC gaggatgagaaaTAACGTGCTGACTTGCACCATATC caaccttcgtaaTTACTATTCCTCCCGTTAAATAATAGATA gaggaaatagtaaTTAACGAAAGGTTGAGATGTTATGTCAGATC CAGTCTGTCAATTCAATCTGCTTCG
<i>recA::tet</i>		F1-F F2-R F3-F F3-R F4-F F4-R	CGTGAAGCAGAATGTGCTTG agtcagcacgtaTTCTCAAATTCGAGTCTTCTTGTG ctcgaattgaagaaTAACGTGCTGACTTGCACCATATC gaaactattttattttaTTACTATTCCTCCCGTTAAATAATAGATA gaggaaatagtaaTAAAAATAAAATAAGTTTCAAATGATACAAAAGG TGTCGGTTGGGAAAATGAGG
<b><i>recA</i> point mutant construction</b>			
	R58C	F1-R F2-F	CAATAATgcaTCCGCGAG CTCGCGGAtgcATTATTG
	K70R	F1-R F2-F	CACAGTTGTacgACCTGAGC GCTCAGGTcgtACAACCTGTG
	E154V	F1-R F2-F	ATGTCGCccaCAATTTCCGC GCGGAAATTGtggGCGACAT
	E154R	F1-R F2-F	ATGTCGCCacgAATTTCCGC GCGGAAATTcgtGGCGACAT
	G155P	F1-R F2-F	ATGTCcggTTCAATTTCCGC GCGGAAATTGAAccgGACAT
	G155R	F1-R F2-F	CCATGTCacgTTCAATTTCCGC GCGGAAATTGAAcgtGACATGG
	D159A	F1-R F2-F	CGACATGCGAcgcTCCCATGTTCG GCGACATGGGAgcgTCGCATGTTCG
	G202I	F1-R F2-F	CCGGGTaatGAACATAACACCG CGGTGTTATGTTcattAACCCGG
	F215Q	F1-R F2-F	GAATAttgTTCAACGCACGCGCG CGGCCGTGCGTTGAAAcaaTATTC
	K241Q K243A	F1-R F2-F	CACGACgttGATttgCGTTTTGTTC GAACAAAACGcaaATCaacGTCGTG
<i>recA::muB</i>		F1-F F1-R F2-R F2-F F3-F F3-R	GTCGTCTCAGGCTGGGAATATTC gcgcgaatatcggaatattcatTCTATTTTTTCTCCTTTATGTTACCACTAC tcaatttcacttcataaccgtaTTAATTACGCAGCAGCGTTGAAATATC gaaaaaatagaATGAATATTTCCGATATTCGCGCAGGAC gcgtaattaataaCGGTTATGAAGTGAAATTGA CGGAAATACTTGTCAAGCTTGCCATC



Information	Primer name	Sequence (5'>3')
<b>Detection of direct repeat sequences</b>		
Cassette Primer	C1	GTACATATTGTCGTTAGAACGCGTAATACGACTCA
Cassette Primer	C2	CGTTAGAACGCGTAATACGACTCACTATAGGGAGA
mini-IS <i>cat</i>	S1	CAGGGTAAAATTTATATCCTTCTTG
mini-IS <i>cat</i>	S2	CCTTCTTGTTTTATGTTTCGG
mini-IS <i>cat</i>	S3	AGAATTGTCAGATAGGCC
mini-IS <i>cat</i>	S4	ATAGGCCTAATGACTGGA
<i>yabC</i>	A-F	GGGCAAACGAGGATCAGATA
	A-R	GCGTCGACTCTTGGAATCGC
<i>queF</i>	B-F	CAGCTGCACACCCTGTTATG
	B-R	CGTGAATTTGCCCCATACTT
<i>tasA</i>	C-F	CGAAGCTACACAGTGGAACG
	C-R	CGAAAAAGTCTCCGCTGTTC
<i>bceB</i>	D-F	TAAAAAGCAAGCACGAGACG
	D-R	GATCCGGCTGAAATGTTTTG
<i>comP</i>	E-F	TGGTCTGAGACAGGGAGACA
	E-R	AAGAATATAGGCAGCAAATGATCTCC
<i>atpH</i>	F-F	AAAGCCGATGAGTGGATCAG
	F-R	GGTTCCAATGCGGACTTTA



付録 5 図. 作製した枯草菌株のゲノム上の構造.

A. *comPd*, *comAd*, *comKd*, *rokd*, *recAd*, *recOd*, *recUd*の遺伝子破壊方法を示す. “Antibiotic resistance gene”は抗生物質耐性遺伝子, “Target gene”は破壊対象の遺伝子を示す.

B. *recA*変異株における変異型*recA*置換株の作製方法. アスタリスク (\*) は*recA*遺伝子上に変異を有することを示す. *emr*は Erythromycin耐性遺伝子.

C. *recA::muB*置換株の作製方法を示す. *tet*は Tetracycline耐性遺伝子を示す.

# 実験に用いた菌株リスト

第2章及び第3章で実験に使用したバクテリアの種名及び遺伝型，プラスミド名のリストを掲載する。

Strain or plasmid	Genotype
<b>B. subtilis strains</b>	
168	<i>trpC2</i>
NBS801	<i>trpC2 trpC2 amyE'::Phyper-spanc IS256Bsu1 tnp IRL-cat-IRR spc::'amyE</i>
NBS2648	<i>trpC2 trpC2 amyE'::Phyper-spanc IS256Bsu2 tnp IRL-cat-IRR spc::'amyE comK::tet</i>
NBS2649	<i>trpC2 trpC2 amyE'::Phyper-spanc IS256Bsu3 tnp IRL-cat-IRR spc::'amyE comP::tet</i>
NBS2650	<i>trpC2 trpC2 amyE'::Phyper-spanc IS256Bsu4 tnp IRL-cat-IRR spc::'amyE comA::tet</i>
NBS2651	<i>trpC2 trpC2 amyE'::Phyper-spanc IS256Bsu5 tnp IRL-cat-IRR spc::'amyE rok::tet</i>
NBS2652	<i>trpC2 trpC2 amyE'::Phyper-spanc IS256Bsu6 tnp IRL-cat-IRR spc::'amyE recA::tet</i>
NBS2653	<i>trpC2 trpC2 amyE'::Phyper-spanc IS256Bsu7 tnp IRL-cat-IRR spc::'amyE recO::erm</i>
NBS2654	<i>trpC2 trpC2 amyE'::Phyper-spanc IS256Bsu8 tnp IRL-cat-IRR spc::'amyE recU::erm</i>
NBS2655	<i>trpC2 trpC2 amyE'::Phyper-spanc IS256Bsu9 tnp IRL-cat-IRR spc::'amyE recA K70R-erm</i>
NBS2656	<i>trpC2 trpC2 amyE'::Phyper-spanc IS256Bsu10 tnp IRL-cat-IRR spc::'amyE recA R58C-erm</i>
NBS2657	<i>trpC2 trpC2 amyE'::Phyper-spanc IS256Bsu11 tnp IRL-cat-IRR spc::'amyE recA E154R-erm</i>
NBS2658	<i>trpC2 trpC2 amyE'::Phyper-spanc IS256Bsu12 tnp IRL-cat-IRR spc::'amyE recA E154V-erm</i>
NBS2659	<i>trpC2 trpC2 amyE'::Phyper-spanc IS256Bsu13 tnp IRL-cat-IRR spc::'amyE recA G155P-erm</i>
NBS2660	<i>trpC2 trpC2 amyE'::Phyper-spanc IS256Bsu14 tnp IRL-cat-IRR spc::'amyE recA G155R-erm</i>
NBS2661	<i>trpC2 trpC2 amyE'::Phyper-spanc IS256Bsu15 tnp IRL-cat-IRR spc::'amyE recA D159A-erm</i>
NBS2662	<i>trpC2 trpC2 amyE'::Phyper-spanc IS256Bsu16 tnp IRL-cat-IRR spc::'amyE recA G202I-erm</i>
NBS2663	<i>trpC2 trpC2 amyE'::Phyper-spanc IS256Bsu17 tnp IRL-cat-IRR spc::'amyE recA F215Q-erm</i>
NBS2664	<i>trpC2 trpC2 amyE'::Phyper-spanc IS256Bsu18 tnp IRL-cat-IRR spc::'amyE recA K241Q K243A-erm</i>
NBS2665	<i>trpC2 trpC2 amyE'::Phyper-spanc IS256Bsu19 tnp IRL-cat-IRR spc::'amyE recA::muB-tet</i>
<b>E.coli strains</b>	
DH110B	<i>F-</i> , <i>mcrA</i> , $\Delta$ ( <i>mrr-hsdRMS-mcrBC</i> ), $\phi$ 80 <i>dlacZ</i> AM15, $\Delta$ <i>lacX74</i> , <i>deoR</i> , <i>recA1</i> , <i>araD139</i> , $\Delta$ ( <i>ara,leu</i> )7697, <i>galU</i> , <i>galK</i> , $\lambda$ -, <i>rpsL</i> , <i>endA1</i> , <i>nupG</i>
<b>plasmid</b>	
pDR111a	<i>bla spc</i>

# 転移頻度測定結果表

第2章, 第3章におけるmini-ISの転移頻度測定結果表を掲載する.

Jumping cat assay strain	Accession ID	TPF (MEAN)	N	SD	P value	Ratio(/WT)
WT (CI)	NBS801	9.336E-06	7	2.900E-06		
WT (LB)	NBS801	1.796E-06	6	1.068E-06	0.0022	0.19237
WT (2xSG)	NBS801	7.490E-06	4	2.717E-06	0.4000	0.80220
WT (2xSG spore)	NBS801	3.161E-06	4	1.491E-06	0.0190	0.33860
<i>comKd</i>	NBS2648	5.630E-07	6	1.752E-07	0.0022	0.06030
<i>comPd</i>	NBS2649	5.917E-07	6	7.668E-07	0.0022	0.06337
<i>comAd</i>	NBS2650	2.912E-07	6	1.383E-07	0.0022	0.03119
<i>rokd</i>	NBS2651	3.672E-06	5	1.155E-06	0.0087	0.39330
<i>recOd</i>	NBS2653	2.400E-05	4	5.354E-06	0.0095	2.57058
<i>recUd</i>	NBS2654	5.150E-05	4	2.505E-05	0.0095	5.51605
<i>recAd</i>	NBS2652	1.203E-08	6	1.337E-08	0.0022	0.00129
<i>recA R58C</i>	NBS2656	1.847E-06	3	2.206E-06	0.0238	0.19779
<i>recA K70R</i>	NBS2655	0.000E+00	3	0.000E+00	0.0238	0.00000
<i>recA E154R</i>	NBS2657	2.674E-05	4	2.039E-05	0.1143	2.86433
<i>recA E154V</i>	NBS2658	3.207E-05	3	2.178E-06	0.0238	3.43459
<i>recA G155R</i>	NBS2660	4.015E-05	4	3.765E-05	0.0381	4.30037
<i>recA G155P</i>	NBS2659	5.500E-06	3	6.144E-06	0.6190	0.58909
<i>recA D159A</i>	NBS2661	1.949E-05	3	1.701E-06	0.0238	2.08787
<i>recA G202I</i>	NBS2662	7.243E-05	4	6.696E-05	0.1143	7.75757
<i>recA F215Q</i>	NBS2663	5.683E-06	4	6.140E-06	0.6571	0.60872
<i>recA K241Q K243N</i>	NBS2664	1.206E-05	3	2.099E-06	0.2024	1.29167
<i>recA::muB</i>	NBS2665	8.690E-06	3	1.369E-06	0.4048	0.93077

付録 7 表 転移頻度測定結果表

mini-IS転移頻度の測定結果表を示す。転移頻度 (TPF) は平均値を示す。「N」は試行回数、「SD」は標準偏差である。「P value」は、並べ替えBrunner-Munzel検定によるCI培地を用いて測定した野生株 (WT(CI)) との検定結果を示す。「Ratio(/WT)」はCI培地を用いて測定した野生株 (WT(CI)) の転移頻度に対する各株の転移頻度の割合を示す。尚、WT(LB), *comKd*, *comPd*, *comAd*, *rokd*, *recAd*の転移頻度は茂木 俊丞氏の測定値である (茂木. 東京農業大学大学院修士論文. 2010.)。ただし、茂木氏の*rokd*の測定値における極端な外れ値は、解析に使用しない (転移頻度"1.46E-05",  $p = 0.0530$ , Smirnov-Grubbs test, この値を除いた同様の外れ値検定では外れ値は生じない)。

# ゲノム構造と高頻度変異領域の関係性を 探るためのシミュレーション解析

ここでは、第4章、第2節で紹介した、ゲノムの構造と変異の生じる座標の関係についてシミュレーションした結果の詳細について記載する。

ゲノムの三次元構造は, Lieberman-Aidenらによる, 立体的なヒルベルト曲線に順ずるフラクタル構造を有することが示された[1,2,3]. 合田徳夫博士の発見である, ヒトゲノム上のSNPsの分布がべき分布であり, かつフラクタル構造を有する時, その原因がゲノム自体の細胞内における構造と, 高頻度変異領域 (Hotspot) とでも言うべき限定された変異の生じやすいゲノム領域に由来するならば, シミュレーションによりその変異の分布傾向とフラクタル性が確認できるはずであると筆者は考えた. ヒトゲノムの細胞内における構造と言えば, クロマチンと呼ばれるタンパク質に巻き取られた構造体が高度に凝集してX字型やY字型などの形状で観察されるイメージが強い. しかし, 分裂期にない細胞では, ゲノムは緩んだ状態で核内に広がっている. Lieberman-Aidenらは, 細胞をホルムアルデヒドで処理することで, 近接するゲノムDNA間を架橋した後, 架橋の起きていない直鎖DNAを切断, ゲノム解析結果に基づきゲノムの展開図を作製, 逆算して立体構造を求めている[2].

## <方法>

そこで, 筆者は古澤満博士, 及び合田徳夫博士の指導の下, 以下の方法でシミュレーションを行った. 16x16のマス目を用意し, ヒルベルト曲線がマス目を通過する順にマス目に番号を振った. この図形をヒルベルト曲線状態のゲノムとし, マス目の番号はゲノム上の座標であると定義した.

1. この図形上に, 任意の大きさの円を乗せ, 円の内部および円周上に存在するマス目の番号を記録した. この試行を10回繰り返した.
2. この図形上に円が複数個ある場合について検証するため, 1) の行程で得られた番号を一つのデータに集積した. ただし, 重複する座標は削除した.
3. 1. 及び2. で得られた数値を各々昇順に並べ, ボックスカウント法により得られた数値からフラクタル性を有するか確認した. 具体的には, 観察するマス目 (解像度: $r$ ) あたりに数字が入っているかいないかを確認し, 数字が入っているマス目の数を数え上げた. これを, マス目の大きさ (解像度: $r$ ) を1, 2, 4, 8, 16と変更した場合についてそれぞれ求めた ( $N(r)$ ). 得られた, 各解像度における数字入りのマス目の値について, それぞれ10を底とす



る対数を取り、グラフにプロットした。同時に、近似直線及び相関係数を求めた。この時得られたグラフが線型に近似された場合はフラクタル性が認められたと判断でき、グラフの傾きがフラクタル次元を示している（解像度によらず同じ図形が同じ数出現することを示しているため）。

4. 2. で得られたデータに関し、座標間距離の頻度を求め、グラフにプロットした後、近似直線及び相関係数を求めた。

5. 上記の試行とランダムな突然変異の場合を比較検討するため、256あるマス目をそれぞれ5分の1の確率で選択し、その座標を記録、3) 及び4) と同様に解析した。具体的には、1から5の乱数を256個同時に発生させ、1が出たマス目の番号を記録した。

なお、円形と接触あるいは交差したマス目では、変異が一回起きるものとした。円形は高頻度変異領域を模しているが、簡便のため、この領域内の全ての座標で変異が生じたものとしている。

#### <結果, 考察>

結果、円の大きさ、数によらず、選択された座標から算出された値からはフラクタル性が見出された。相関係数 ( $R^2$ ) はほとんどの場合0.99より大であった。マス目の一辺の大きさを1とした時の円の大きさを求め、そのフラクタル次元の相関性を調べると、両者は正の相関関係を有することが分かった ( $R^2=0.84$ )。また、円を10個座標の上に置いた場合はフラクタル次元が最大 (0.9259,  $R^2=0.9985$ ) となった。この値は、ヒトゲノムのSNPsに置けるフラクタル性解析の結果とほぼ同じである (HapMap: 0.96115, iJGVD (Touhoku): 0.96176, 1000 Genomes: 0.99476)。また、円を10個座標の上に置いた場合は選択されている場合について座標間距離の頻度を算出し、分布の状態を調べると、べき分布であることが分かった ( $R^2=0.98509$ )。この点も、ヒトゲノムのSNPsより得られた結果と共通している。

一方、ランダムに突然変異を発生した場合は、円でマス目を選択した場合とは異なり、ボックスカウント法により求めたフラクタル次元は約0.38であった。また、変異間距離の頻度のドットプロットに対する近似直線の傾きを、円によるセル選択時とランダムなセル選択

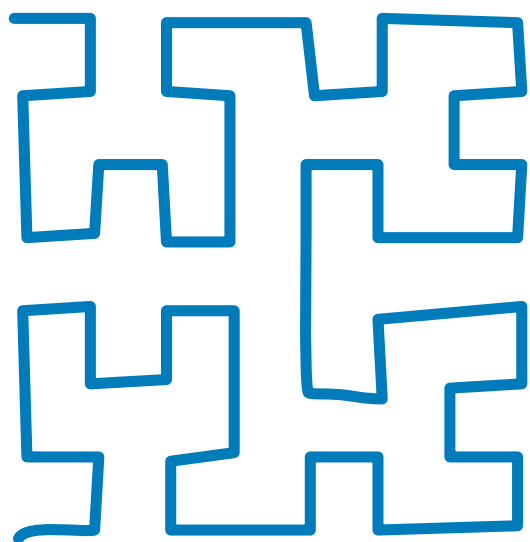
時で比較すると、前者の近似直線の係数が-3.75であるのに対し ( $R^2=0.99$ , 付録 8 図3 B, “All”), 後者は-0.57であり ( $R^2=0.67$ , 付録 8 図4 B), ランダムにセルを選択した場合の近似直線の傾きの方が小さい。ランダムに座標が選択された場合, 座標間距離の頻度は一様分布となるはずであるから, 傾きがランダムの場合で小さいのは一様分布に近い分布であることを意味すると考えられる。まとめると, ランダムにセルを選択した場合において, フラクタル次元が実測値よりも低いこと, 両対数グラフにおける線形近似直線の傾きが円形で選択した場合と比べて小さく, その相関係数も低いことから, セルをランダムに選択した場合のシミュレーション結果は実際の結果を反映していないことが示唆された。これらより, ランダムにヒルベルト曲線上の座標を選択した場合よりも, 円で選択した場合の方が, ヒトゲノムのSNPsの解析結果を反映している可能性が高いと考えられる。

以上のシミュレーション結果から, ヒルベルト曲線に座標を割り当て, 円形により選択された座標の性質について調べると, その分布はべき分布であり, 分布の幾何的な構造にはフラクタル性があることが明らかとなった。SNPsの分布がべき分布であり, かつフラクタル性が認められる要因は, ヒトゲノムがヒルベルト曲線に順ずる構造を有すること及び以上のシミュレーション結果を考慮すると, ゲノム上にその物理的な構造に依拠した突然変異の発生しやすい領域 (Hotspot) が存在するためであると考えられる。

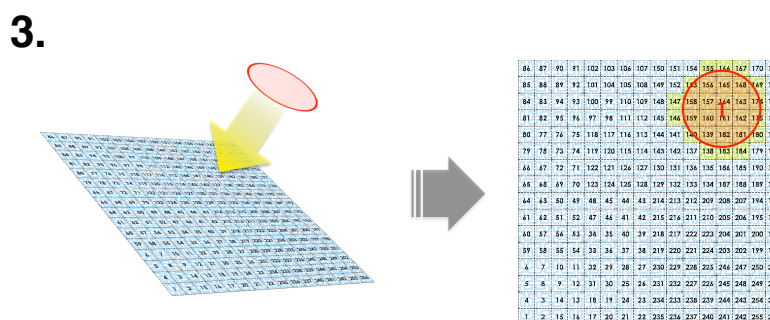
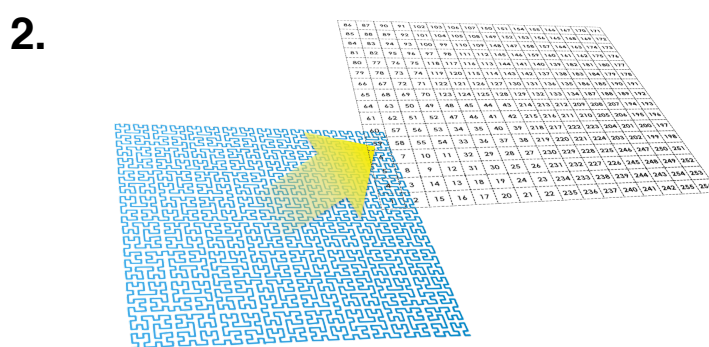
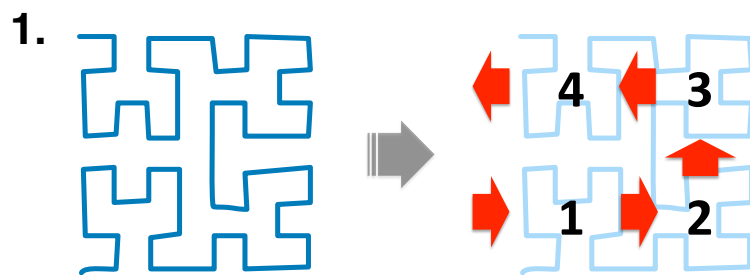
## 参考文献

1. Hilbert D. Ueber die stetige Abbildung einer Linie auf ein Flaechenstueck. *Math Ann.* 1891;38(3):459-460.
2. Lieberman-Aiden E, van Berkum NL, Williams L, et al. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science.* 2009;326(October):289-293.
3. Mirny L a. The fractal globule as a model of chromatin architecture in the cell. *Chromosom Res.* 2011;19:37-51.

**A**



**B**

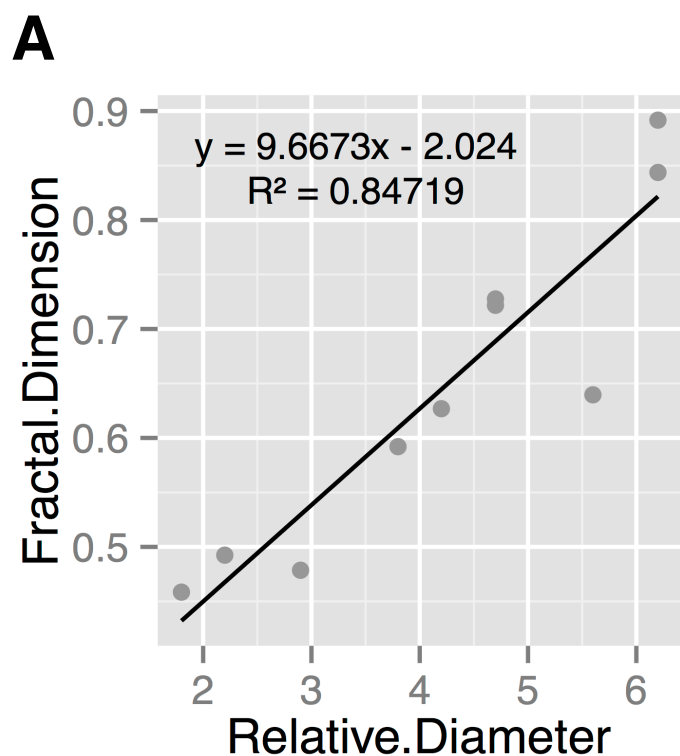


付録 8 図1. 突然変異分布状況のシミュレーション

A. ヒルベルト曲線及び、シミュレーションの際の数え上げの単位となる構造を示す。

B. シミュレーションのマトリクス作製及びシミュレーションの手順を示す. 1) マトリクスの各セルへの番号付のルール. 2) 16x16のマトリクスとなるようにヒルベルト曲線を作製し、マトリクスを重ねて番号を振った. 3) 作製したマトリクスに任意の大きさの円を乗せ、円に接するか内包されたセルの番号を記録した。





**B**

Circle No.	Relative Diameter	Fractal Dimension	R <sup>2</sup>
1	2.9	0.4785	0.93841
2	4.7	0.7275	0.98112
3	6.2	0.8917	0.99541
4	2.2	0.4925	0.93394
5	4.2	0.6269	0.99036
6	5.6	0.6396	0.97529
7	4.7	0.7218	0.99107
8	3.8	0.5919	0.99013
9	1.8	0.4585	0.92455
10	6.2	0.8437	0.99786
All		0.9259	0.99856

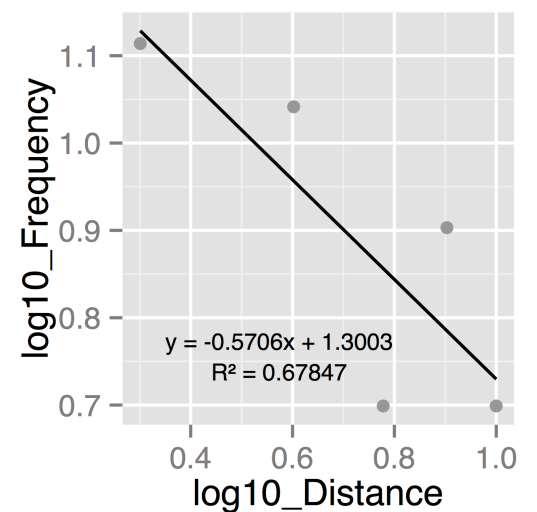
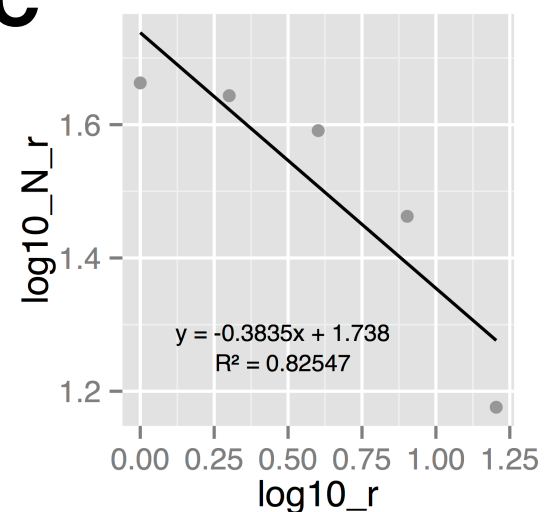
**付録 8 図3. マトリクスに配置した円の大きさとフラクタル次元の関係**

A. 10回行ったシミュレーションに用いた円の（セルの一边を1とした時の）大きさを求め、得られたフラクタル次元とともにプロットしたグラフ。縦軸（Fractal Dimension）はフラクタル次元、横軸（Relative Diameter）は円の直径を示す。上端左の数式はデータに対する線型近似直線、 $R^2$ は相関係数。

A. 10回行ったシミュレーションに用いた円の相対的な大きさ（Relative Diameter）、それにより得られたフラクタル次元（Fractal Dimension）、及びフラクタル次元を求める際に得られた近似直線の相関係数（ $R^2$ ）を示す。「Circle No.」はシミュレーション番号と対応しているが、「All」はすべてのデータを重ね合わせた場合の値を示すため、円の大きさは記載していない。

**A**

86	87	90	91	102	103	106	107	150	151	154	155	166	167	170	171
85	88	89	92	101	104	105	108	149	152	153	156	165	168	169	172
84	83	94	93	100	99	110	109	148	147	158	157	164	163	174	173
81	82	95	96	97	98	111	112	145	146	159	160	161	162	175	176
80	77	76	75	118	117	116	113	144	141	140	139	182	181	180	177
79	78	73	74	119	120	115	114	143	142	137	138	183	184	179	178
66	67	72	71	122	121	126	127	130	131	136	135	186	185	190	191
65	68	69	70	123	124	125	128	129	132	133	134	187	188	189	192
64	63	50	49	48	45	44	43	214	213	212	209	208	207	194	193
61	62	51	52	47	46	41	42	215	216	211	210	205	206	195	196
60	57	56	53	34	35	40	39	218	217	222	223	204	201	200	197
59	58	55	54	33	36	37	38	219	220	221	224	203	202	199	198
6	7	10	11	32	29	28	27	230	229	228	225	246	247	250	251
5	8	9	12	31	30	25	26	231	232	227	226	245	248	249	252
4	3	14	13	18	19	24	23	234	233	238	239	244	243	254	253
1	2	15	16	17	20	21	22	235	236	237	240	241	242	255	256

**B****C**

#### 付録 8 図4. セルをランダムに選択した場合のシミュレーション結果

- A. シミュレーション解析に用いた図形. 黄色のセルはランダムに選択されている.
- B. A図のシミュレーションにより得られた座標間距離の頻度を求め、頻度と距離について対数をとってプロットしたグラフ (log10\_Frequency: log10(頻度), log10\_Distance: log10(座標間距離)). 左下の数式はデータに対する線型近似直線,  $R^2$ は相関係数.
- C. A図のシミュレーションにより得られた座標の数をボックスカウント法で求め、解像度ごとにプロットした図 (log10\_N\_r: log10(各解像度における座標の数), log10\_r: log10(解像度)). 左下の数式はデータに対する線型近似直線,  $R^2$ は相関係数. グラフの傾きの絶対値 (0.9259) はフラクタル次元を意味する.

# 培地, 試薬, プロトコール

本研究において使用した基礎的な試薬, 培地, および実験プロトコールを掲載する.

## 培地および試薬

### 培地

#### LB (800 ml)

Tryptone peptone (Difco Laboratories)	8 g
Yeast extract (Difco Laboratories)	4 g
NaCl (Difco Laboratories)	4 g
細菌用粉末寒天 (日水製薬)	12 g
H <sub>2</sub> O	800 ml

上記の材料を1 L三角フラスコに混ぜ、オートクレーブで121°C, 20分滅菌する. 50°C程度まで冷まし, 各抗生物質 (作り方: 別紙参照) を加え, 良く攪拌した後, バーナーを焚いてプレートに入れる.

#### 2xSG (100 ml)

* Nutrient Broth (Difco Laboratories)	1.6 g
* KCl (柳島製薬株式会社)	0.1 g
* 細菌用粉末寒天 (日水製薬株式会社)	2 g
* H <sub>2</sub> O	100 ml

1M MgSO <sub>4</sub>	100 μl
1M Ca(NO <sub>3</sub> ) <sub>2</sub>	100 μl
1mM FeSO <sub>4</sub>	100 μl
10mM MnCl <sub>2</sub>	100 μl
50% glucose	200 μl

(\*) の材料を試薬ビンに混ぜ, オートクレーブで121°C, 20分滅菌する. オートクレーブ後, 残りの試薬を加え, 遮光する.



10×MM <Minimal Medium> (250 ml)

K <sub>2</sub> HPO <sub>4</sub> (関東化学)	35 g
KH <sub>2</sub> PO <sub>4</sub> (関東化学)	15 g
(NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> (関東化学)	5 g
Trisodium Citrate Dihydrate (関東化学)	2.5 g

ビーカーに約100 mlの水とスターラーを入れておき, 上記の試薬を加えて攪拌しながら溶解する. メスシリンダーに洗いこみながら移して250 mlにfill upし, 試薬ビンに移してオートクレーブで121°C・20分滅菌する. 室温で保存し, 使用時に10倍希釈 (1×MM) にして再度オートクレーブ滅菌する.

CI (5 ml)

1×MM	5 ml
50% Glucose	50 μl
1M MgSO <sub>4</sub>	25 μl
5% CAA	20 μl
2.5 mg/ml Tryptophan	100 μl

用事調整.

CII (5 ml)

1×MM	5 ml
50% Glucose	50 μl
1M MgSO <sub>4</sub>	25 μl
5% CAA	10 μl
2.5 mg/ml Tryptophan	10 μl

用事調整.

## 試薬類

### Spectinomycin (100 mg/ml)

スペクチノマイシン (ナカライテスク) 1 g

H<sub>2</sub>O 10 ml

使用済みCORNIGチューブ (オートクレーブ済み) に計り取り水を加え, Vortex溶解後  
フィルター滅菌 (0.45 $\mu$ m) し1.5 ml チューブに分注して-30°Cに保存.

### Chloramphenicol (5 mg/ml)

クロラムフェニコール (ナカライテスク) 0.05 g

99.5% EtOH 10 ml

TPP 15 mlチューブに計り取り, 99.5% EtOHを加え, Vortex溶解後-30°Cに保存.

### Erythromycin (1 mg/ml)

エリスロマイシン (ナカライテスク) 0.01 g

99.5% EtOH 10 ml

TPP 15 mlチューブに計り取り, 99.5% EtOHを加え, Vortex溶解後-30°Cに保存.

### Rifampicin (Rifampin) (5 mg/ml)

Rifampicin (SIGMA) 0.005 g

DMSO 1 ml

1.5 mlエッペンチューブに計り取り, DMSOを加え, Vortex溶解後使用. 用事調整.

### RNase (10 mg/ml)

Ribonuclease A (SIGMA)	10 mg
1M Tris-HCl (pH7.5)	10 $\mu$ l
1M NaCl	15 $\mu$ l
H <sub>2</sub> O	975 $\mu$ l

よく混ぜて溶かした後, 100°C, 15min incubateし室温でゆっくり常温に戻す.

### Lysozyme (10 mg/ml)

Lysozyme	10 mg
10mM Tris-HCl (pH8.0)	1 ml

用事調整で溶解して使用する.

### 10% SDS (100ml)

SDS (カライクス)	10 g
-------------	------

試薬ビンをおートクレーブにかけ, 乾燥させておく. 適当な大きさのビーカーにSDS を10g入れ, 加温しながら 水に溶かす (高温不可・50-60°Cまで, Volumeが出やすいため注意). マグネチックスターラーを使用する. 溶解後, 洗い込みながら, 100mlにfill up・フィルター滅菌で滅菌乾燥した試薬ビンに入れる. なお, Sodium dodecyl sulfate は Sodium lauryl sulfateに同じ.

### 3M NaOAc (pH5.2) (100ml)

酢酸ナトリウム (関東化学)	24.61g
----------------	--------

酢酸 (関東化学)

酢酸ナトリウムをビーカーに加えた適当量の水でマグネチックスターラーを回しながら溶かし, 酢酸でpHを調整した後メスシリンダーで100mlにfill up, 試薬ビンに移す.

### 70% Ethanol

200 mlの青フタガラス瓶に60 mlの水を入れてオートクレーブする。オートクレーブ後に99% EtOHを140 ml加えて, -30°C保存する。

### 80% Glycerol

グリセリン（関東化学）に水を加えてオートクレーブする。

### 1M MgSO<sub>4</sub>

適当な大きさのビーカーにMgSO<sub>4</sub>を量り入れ, fill upすることを考えて, 適当量の水を入れ, 加温しながらマグネチックスターラーを回して溶かした後, メスシリンダーで100 mlにfill upし, 試薬ビンに入れてオートクレーブする。

### 50%glucose (100 ml)

glucose（関東化学） 50 g

H<sub>2</sub>O X ml

スターラーを入れた100 mlのビーカーに50 ml程度の水を入れる。量り取ったglucoseをビーカーに少量ずつ入れ, 全て溶かす（熱を加えるとより溶け易くなる）全て解けたら, 100 mlのメスシリンダーに注ぎ, ビーカーを2, 3度洗いこむ。100 mlまでfill upし, パラフィルムで蓋をし, よく混ぜる。試薬ビンに分注し, オートクレーブで121°C・15 min滅菌する。

### 5% Casamino acid

Bacto™ VITAMIN ASSAY CASAMINO ACIDS (DIFCO) 5 g

試薬ビンをオートクレーブにかけ, 乾燥させておく。適当な大きさのビーカーにCASAMINO ACIDSを5 g入れ, 加温しながら（高温不可・50-60°Cまで）水に溶かす。マグネチックスターラーを使用する。溶解後, 洗い込みながら, 100 mlにfill up・フィルター滅菌で滅菌乾燥した試薬ビンに入れる。

## 2.5 mg/ml L-Tryptophan

L-Tryptophan (関東化学) 250 mg

試薬ビンオートクレーブにかけ、乾燥させておく。適当な大きさのビーカーにL-Tryptophanを入れ、加温しながら（高温不可・50-60°Cまで）水に溶かす。マグネチックスターラーを使用する。溶解後、洗い込みながら、100 mlにfill up・フィルター滅菌で滅菌乾燥した試薬ビンに入れ、試薬ビンはアルミホイルで遮光し、4°C保存。

## TES buffer (100 ml)

Sucrose 25 g (final conc. 25%)

1M Tris- HCl (pH7.6) 5 ml (final conc. 50 mM)

0.5M EDTA (pH8.0) 4 ml (final conc. 20 mM)

ビーカー中に適当量の水（約50 ml）を入れて加熱しながらマグネチックスターラーを回し、Sucroseを溶かす。溶解後、洗い込みながら100 mlのメスシリンダーに入れ、Tris-HClとEDTAを加えFill up後オートクレーブで滅菌。

## 0.5M EDTA (pH8.0) (500 ml)

EDTA-3Na (FW=412.24) (カライケス) 103.06 g

ビーカー中の適当量の水（約350 ml）にマグネチックスターラーを回しながら、10 N NaOH（約5 ml）でpHを調整する。溶解した後メスシリンダーで500 mlにfill up, 試薬ビンに移してオートクレーブする。

## 1M Tris-HCl (100ml)

TRIZMA HCl (SIGMA) 0.444 g

TRIZMA Base (SIGMA) 0.265 g

25°C (pH8.0) のもので水に溶かし、メスシリンダーで100 mlにfill upし、試薬ビンに入れてオートクレーブする。

## 50×TAE (1000ml)

Tris Base (SIGMA)	242 g
Acetic acid (関東化学)	57.1 ml
0.5M EDTA (調製法：別紙参照)	100 ml(pH8.0)

1Lのビーカーに500-600 mlのmilliQ, Tris-Baseを入れてスターラーで攪拌. 吸熱反応のため, 温めると溶けやすくなる. 酢酸 (発熱反応) と, EDTAを加えて1000 mlにfill upする. オートクレーブはかけない.

## 実験プロトコール 1

### 枯草菌からのゲノムDNA抽出-1

PCR templateに用いるDNAを調整するためのプロトコール. 全行程は1-2日.

セレクション用の薬剤を添加したLB培地に菌を爪楊枝で植菌し, 前培養する.

↓

上清を除去した後, 飽和ヨウ化カリウムを100  $\mu$ l加える.

↓

菌体を爪楊枝で取り, Vortexする.

↓

37°Cでインキュベートする(O/N可).

↓

1 mlの無菌水を加え, 転倒攪拌する.

↓

遠心 (4°C, 15,000rpm, 1分)

↓

上清を除去する. 保存する場合は-80°Cへ入れる.

↓

以下のものを加えてよく懸濁する.

- TESバッファー 100  $\mu$ l
- Lysozyme 5  $\mu$ l
- RNase 1  $\mu$ l

↓

37°Cで30分間インキュベートする(O/N可).

↓

10% SDSを5 µl加えて穏やかに転倒混和する.

↓

フェノール/クロロホルムを100 µl加え, Vortexする.

↓

遠心 (4°C , 15000rpm , 5分)

↓

DNA溶液に以下のものを加えてしっかりと混和する.

- ・ DNA溶液の1/10量の酢酸ナトリウム
- ・ DNA溶液の2.5倍量の特級99.5%エタノール

↓

遠心 (4°C , 15000rpm , 5分)

↓

-80°Cで10分間冷却する.

↓

デカントで上清を捨て, -30°Cの70%エタノールを500µl加え, 壁面を軽く洗う.

↓

遠心 (4°C , 15000rpm , 5分)

↓

上清をデカントとピペットマンで取り除く.



↓

減圧乾燥

↓

滅菌蒸留水を20  $\mu$ l添加し, Vortexする.

↓

希釈(x20 $\leq$ )してPCRに用いる.

## 実験プロトコール 2

### 枯草菌からのゲノムDNA抽出-2

サザンブロッティング, 次世代シーケンサー用など, DNAそのものが解析対象となる実験に適したプロトコール. 全行程は3日を要す.

LBプレーンプレートにフリーズストックの菌体を3段階に塗り広げ, 25°Cで12-16 h前培養する.

↓

LB培地5 mlを含むL字管 に植菌し, O.D.600=0.1に植菌, 定常期に入るまで37°Cで振盪培養する.

↓

培養液を1.5 mlエッペンチューブ数本にデカントで移す.

↓

遠心 (4°C , 15,000rpm , 1分)

↓

上清を除去した後, 飽和ヨウ化カリウムを500  $\mu$ l加える.

↓

37°Cで30分間インキュベート.

↓

1 mlの無菌水を加え, 転倒攪拌する.

↓

遠心 (4°C , 15,000rpm , 1分)

↓

上清をアスピレーターで除去して, 以下のものを加えてよく懸濁する.

- TESバッファー 500  $\mu$ l
- Lysozyme 20  $\mu$ l

↓

37°Cで30分間インキュベート.

↓

10% SDSを20  $\mu$ l加えて穏やかに転倒混和する.

↓

フェノール/クロロホルムを500  $\mu$ l加え, 完全に混ざるまで転倒攪拌する.

↓

遠心 (4°C , 15000rpm , 10分)

↓

先端を切り取ったチップをつけたピペットマンを用いて, 新しい1.5 mlエッペンチューブに上清のみを移す (注1) .

↓

フェノール/クロロホルムを上清と等量加え, 完全に混ざるまで転倒攪拌する.

↓

遠心 (4°C , 15000rpm , 10分)

↓

先端を切り取ったチップをつけたピペットマンを用いて, 新しい1.5 mlエッペンチューブに上清のみを移す (注2) .

↓

クロロホルムを上清と等量加え, 完全に混ざるまで転倒攪拌する.

↓

遠心 (4°C , 15000rpm , 10分)

↓

DNA溶液に以下のものを加えてしっかりと転倒攪拌する.

- ・ DNA溶液の 1 /10量の酢酸ナトリウム
- ・ DNA溶液の2.5倍量の特級99.5%エタノール

↓

-80°Cで10分間冷却する.

↓

遠心 (4°C , 15000rpm , 10分)

↓

デカントで上清を捨て, -30°Cの70%エタノールを500μl加え, 壁面を軽く洗う.

↓

遠心 (4°C , 15000rpm , 5分)

↓

上清をデカントとピペットマンで取り除く.

↓

減圧乾燥

↓

滅菌蒸留水を200 μl加える.

- ↓
- RNaseを1  $\mu$ l加えて37°CでO/Nさせる.
- ↓
- フェノール/クロロホルムを200  $\mu$ l加え, 完全に混ざるまで転倒攪拌する.
- ↓
- 遠心 (4°C , 15000rpm , 5分)
- ↓
- 先端を切り取ったチップをつけたピペットマンを用いて, 新しい1.5 mlエッペンチューブに上清のみを移す.
- ↓
- DNA溶液に以下のものを加えてしっかりと混和する.
- ・ DNA溶液の1/10量の酢酸ナトリウム
  - ・ DNA溶液の2.5倍量の特級99.5%エタノール
- ↓
- 80°Cで10分間冷却する.
- ↓
- 遠心 (4°C , 15000rpm , 10分)
- ↓
- デカントで上清を捨て, -30°Cの70%エタノールを500 $\mu$ l加え, 壁面を軽く洗う.
- ↓
- 遠心 (4°C , 15000rpm , 5分)
- ↓
- 上清をデカントとピペットマンで取り除く.

↓

減圧乾燥

↓

滅菌蒸留水またはTEを適量加える。

↓

電気泳動により、収量を確認する。また、RNAが混ざっていないか確認する。

↓

O.D.260およびO.D.280の吸光度を測定し、電気泳動で確認した収量と一致するか確認する。

注1. 中間層の白い塊、下層はなるべく吸わないようにすること。ただし、ゲノムDNAは粘性が強く、吸いにくい上に、中間層・下層も一緒に吸いやすい。そのため、もし上清のみを吸いにくい場合は中間層ごと回収した後、以下の操作で中間層を除去する。

- ・ 中間層ごと、まとめて新しいエッペンチューブに移す

↓

- ・ 遠心 (4°C, 15000rpm, 10分)

↓

- ・ sup

注2. 以後、白い塊が出現しなくなるまで、フェノール/クロロホルムによる精製操作を繰り返す

## 実験プロトコール 3

### 枯草菌コンピテントセルの調整

枯草菌を、形質転換効率が上昇するような条件で培養することにより、コンピテントセルを作製する。

凍結保存された枯草菌168株を、LB plain plateに画線塗抹し、25°CでO/Nする(12~16 h).

↓

L字管にCI液体培地とCII液体培地を各5 ml調整する。

↓

CI液体培地にO.D.600 = 0.05となるように菌体を加え、よくvortexする。

↓

37°Cに設定した恒温槽でL字管を振盪培養する。培養しながら、30分おきにBactomonitorで吸光度を測定し(O.D.600), 片対数グラフにプロットしていく。

↓

吸光度測定値の増加率が低下してきたら、無菌の2 mlエッペンチューブに、培養液を分注する。

O.D.600  $\geq$  0.8なら2.5ml, O.D.600 < 0.8なら5mlの培養液を使用する。

↓

遠心 (25°C , 8,000rpm , 1分)

↓

上澄をピペットマンで除去する。形質転換能のある菌は浮き易いため、上澄は少し残す。

↓

CII培養液で菌体を溶解し，CII液体培地5 mlに加えてvortexし，吸光度（O.D.600）を測定する．

↓

37°Cに設定した恒温槽でL字管を30分間振盪培養する．

↓

吸光度（O.D.600）を測定し，値がCII移行時よりも上昇していることを確認する．

↓

小試験管に200  $\mu$ lの培養液と1  $\mu$ g以下の形質転換用DNAを加え，37°Cで1時間，振盪培養する（注1）．

残った培養液は700  $\mu$ lのエッペンチューブに200  $\mu$ lずつ培養液を分注し，-80°Cで保存する．

↓

振盪培養した培養液を選択プレートに全量塗抹し，37°CでO/Nする．

↓

生えてきたコロニーを数え，形質転換効率を算出する．PCR断片なら数十個，プラスミドやゲノムDNAなら数百個のコロニーが生じる．

注1. この行程でDNAの濃度を0.01  $\mu$ g, 0.1  $\mu$ g, 1.0  $\mu$ g, 10.0  $\mu$ gとした時の形質転換体の数を記録する．形質転換効率はコンピテントセルとDNAの出会う確率に依存するため，添加するDNAの量に比例して形質転換体が生じるはずである．もしこのようにならない場合は異常が起きている可能性があるため，サンプルは使用しないかよく原因を調べてみる必要がある．



## 実験プロトコール 4

### カラムによるPCR断片の精製

Recombinant-PCRを行う際やシーケンス反応を行う際など，PCRのテンプレートとなる断片が入った溶液に，以前の反応に用いたPrimerが残っていると非特異増幅の原因となる．そのため，エタノール沈澱やカラムにより，PCR反応液からPrimerを除去する必要がある．ここでは，PROMEGA社のWizard® SV Gel and PCR Clean-UP Systemを用いたPCR断片の精製方法を紹介する．

Wizard® SV MinicolumnをCollection tubeにセットする．

↓

PCR産物に，PCR反応液と等量のMembrane Binding Solutionを加え，溶液が均一になるまでよくピペッティングする．

↓

Wizard® SV Minicolumnのフィルタ上に，直接Binding Solutionを加えたPCR反応液を乗せ，1分間静置する．

↓

遠心 (4°C , 15,000rpm , 1分)

↓

Collection tubeから，濾過された廃液を除去し，再度Wizard® SV Minicolumnをセットする．

↓

500 µlのMembrane Wash Solution (注1) をWizard® SV Minicolumnに加える．

↓

遠心 (4°C , 15,000rpm , 1分)

↓

Wizard® SV Minicolumnを新しいCollection tubeにセットする.

↓

10分間, 減圧乾燥する (注2) .

↓

Wizard® SV Minicolumnのフィルタ上に, 直接Elution Buffer又は滅菌水を乗せ, 1分間静置する (注3) .

↓

遠心 (4°C , 15,000rpm , 1分)

↓

アガロースゲル電気泳動により, サンプルの精製を確認する.

注1. Membrane Wash Solutionに95%エタノールが加えられているか確認する.

注2. シークエンス反応に用いるサンプルの場合, ここでWash Solutionに含まれるエタノールを揮発させないと, キャピラリーシークエンス時にノイズが出てしまう.

注3. 溶出させる溶液量が50 µlを切ると収率が低下する.

## 実験プロトコール 5

### アガロースゲルによるDNA断片の精製

Recombinant PCRの際, 非特異的な断片が多く混入していると, 目的としている断片どうしのannealing効率が低下する恐れがある. アガロースゲル電気泳動でPCR反応液を泳動し, 目的の断片のみを切り出すことで, PCR断片を精製することが可能である. ここでは, GE Healthcare社のillustra™ GFX™ PCR DNA and Gel Band Purification Kitを用いたDNA断片の精製方法を紹介する.

Lonza SeaKem® GTG® Agaroseを用い, 電気泳動用の1% アガロースゲルを作製する.

↓

PCR反応液を10~20 µlと10 x loading bufferを混ぜ, 作出したゲルを用いて電気泳動する (注1) .

↓

ゲルをエチジウムブロマイドで染色した後, トランスイルミネータを長波に設定して泳動したバンドを確認する.

↓

カミソリで精製したいDNAのバンドを切出し, トランスイルミネータのスイッチを切る.

↓

切り出したアガロースゲルを小口に切り, 1.5 ml エッペンチューブに全て移す.

↓

精密秤で, 切り出したゲルの重さを測定する (注2) .

↓

ゲルと等量のCapture buffer type3.を加え, vortexする.

↓

60°Cで加温し, ゲルを完全に溶解させる. 適宜vortexし, 溶液を均一にする.

↓

illustra™ GFX™ MicroSpin™ columnをCollection tubeにセットする.

↓

illustra™ GFX™ MicroSpin™ columnのフィルタ上に, 溶解したゲルを直接乗せ, 1分間静置する.

↓

遠心 (4°C , 15,000rpm , 1分)

↓

Collection tubeから, 濾過された廃液を除去し, 再度illustra™ GFX™ MicroSpin™ columnをセットする.

↓

500 µlのWash buffer type1 (注3) をillustra™ GFX™ MicroSpin™ columnに加える.

↓

遠心 (4°C , 15,000rpm , 1分)

↓

illustra™ GFX™ MicroSpin™ columnを新しいCollection tubeにセットする.

↓

10分間, 減圧乾燥する (注4) .

↓

illustra™ GFX™ MicroSpin™ columnのフィルタ上に, 直接Elution Buffer type4. (注5) を乗せ, 1分間静置する (注6) .

↓

遠心 (4°C , 15,000rpm , 1分)

↓

アガロースゲル電気泳動により, サンプルの精製を確認する.

注1. 精製したい断片がゲルの中央に来るよう, 電気泳動時間は適宜調節する.

注2. 60°Cで加温する際, ゲルを細かく切っておかないと溶けにくくなる.

注3. Membrane Wash Solutionに95%エタノールが加えられているか確認する.

注4. シークエンス反応に用いるサンプルの場合, ここでWash Solutionに含まれるエタノールを揮発させないと, キャピラリーシークエンス時にノイズが出てしまう.

注5. Elution Buffer type4.は保存用及び他の実験用のサンプルに用いる. Elution Buffer type6.はシークエンス反応用.

注6. 溶出させる溶液量は10-50  $\mu$ lだが, 50  $\mu$ lでの回収率が最も高い.

## 実験プロトコール 6

### 胞子形成率測定プロトコール

枯草菌の胞子形成率を測定するためのプロトコール。胞子形成培地（2xSG培地）で振盪培養した後，培養液を一部採って加熱し，胞子を形成していない菌を殺菌する。殺菌処理した培養液の生菌数と，未処理の培養液の生菌数を比較する事により，胞子形成率を培養液あたりの胞子の存在頻度として求める事ができる。全行程は3日。

(用意するもの)

- L字管 x3
- 小試験管 x36
- LB plane plate x24
- 100 ml LB liquid
- 100 ml 2xSG liquid
- heat block

#### 1. 前培養~本培養

菌体をLB plane plateに25°Cで前培養した。

↓

L字管に2xSG培地を5 ml分注し，O.D.600=0.1となるように植菌した。

↓

恒温槽で振盪培養した。(37°C, 48rpm, 24 h)

↓

培養液を1.5 ml test tubeに1 mlずつ分注した。

## 2. 胞子液のプレーティング

ヒートブロックを用いて培養液を入れた1.5 ml test tubeを80°Cで10 min加熱した（注1）。

↓

加温した培養液を約25 °Cの水につけ, 冷却した。

↓

glass tubeを用いて培養液を $10^6$ までLB液体培地で希釈した。希釈はLB液体培地900  $\mu$ lに培養液100  $\mu$ lを加えて行い, 各段階においてよくVortexした（注2）。

↓

$10^5$ と $10^6$ の培養液を100  $\mu$ lずつ, 2枚のLB plate plateにプレーティングした。

↓

プレーティングしたプレートを37°CでO/Nした。

↓

コロニー数を数え, CFUを算出した（注3）。

## 3. 培養液のプレーティング

glass tubeを用いて培養液を $10^6$ までLB液体培地で希釈した。希釈はLB液体培地900  $\mu$ lに培養液100  $\mu$ lを加えて行い, 各段階においてよくVortexした。

↓

$10^5$ と $10^6$ の培養液を100  $\mu$ lずつ, 2枚のLB plate plateにプレーティングした。

↓

プレーティングしたプレートを37°CでO/Nした。

↓

コロニー数を数え, CFUを算出した.

#### 4. 孢子形成率算出

孢子形成率を, 全細胞のCFUに対する孢子のCFUの割合として算出した.

注1. 80°Cで加温する際は, 容器が密閉されるものを使用しないと, 原液の濃度が変わってしまうため, 実際よりも孢子形成率が高く見積もられてしまう. ガラスは多糖と親和性が高いので使用しない方が良い (孢子の周囲は多糖類の幕なので) .

注2. 希釈に用いるLB培地はガラスピペットを使用して採る. このとき, 1 mlのLB培地を取り, 900  $\mu$ lを排出すると誤差が小さい. 同様にして, 培養液はP-200のピペットマンを用い, リバースピペッティングで採液する (P-200のピペットマンで90  $\mu$ l, P-20のピペットマンで10  $\mu$ lの培養液を加えても良い) .

注3.  $10^6$ ,  $10^5$ のプレートより算出したCFUの平均値を比較する. 両者の値が著しく異なる場合 (10倍以上の測定誤差等) は希釈倍率の低い $10^5$ より求めた値を採用する. 両者の測定値の差が誤差範囲であれば, これらの値から平均値を算出し, CFUとする.



