

日本生物学オリンピック本選2023

静岡大会(静岡)

大問 1



2023年8月18日(金)

制限時間 75分

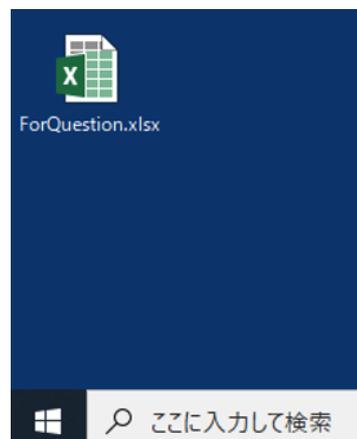
諸注意

1. 机には問題冊子 1 部(12 ページ)と解答用紙 2 枚、グラフ用紙 1 枚が配布されている。
2. 本試験にはコンピュータを用いた情報処理が含まれており、問題に応じて答案を解答用紙に記入する形式と答案を解答ファイルに記入する形式がある。試験中にコンピュータ端末で使用可能なソフトはWord、Excel、メモ帳と指定の解析ソフトのみとする。インターネットを利用するソフト(Chromeなど)は、使用しないこと。
3. 試験中にコンピュータ端末の不具合が発生した場合、作業中のデータが失われることがある。作業中のファイルをこまめに保存すること。万が一、コンピュータ端末に不具合があった場合には挙手すること。
4. 指示があるまでこのページの内容をよく読んでおくこと。試験開始の合図があるまで机の上の機材には触れないこと。
5. 解答開始の合図の後、すべての解答用紙とグラフ用紙、解答ファイルに本選受験番号と氏名を記入すること。
6. 試験の途中で体調が悪くなった場合、トイレに行きたい場合、水分補給をしたい場合、その他やむを得ず途中退出を希望する場合は挙手をして試験監督に知らせること。
7. 問題冊子にメモを取っても構わない。試験終了後、問題冊子は持ち帰ること。
8. 最初に問題冊子の全てのページに目を通してから、解答をはじめること。
9. 机の上は既に準備されているものの他に、ペンケース、筆記用具、時計、ハンカチ、ティッシュ、目薬以外のものはおかないこと。

【重要】試験開始前の準備

解析1の準備

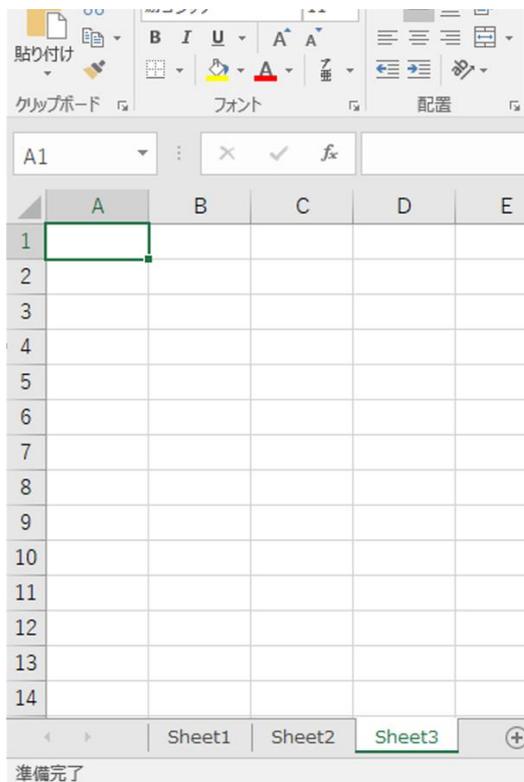
最初に、PCのデスクトップ上に計算用エクセルファイル「ForQuestion.xlsx」があることを確認すること(ただし、試験開始の合図があるまでファイルを開いてはならない)。また、机上に定規があるか確認すること。ない場合は、挙手にて試験監督に知らせること。



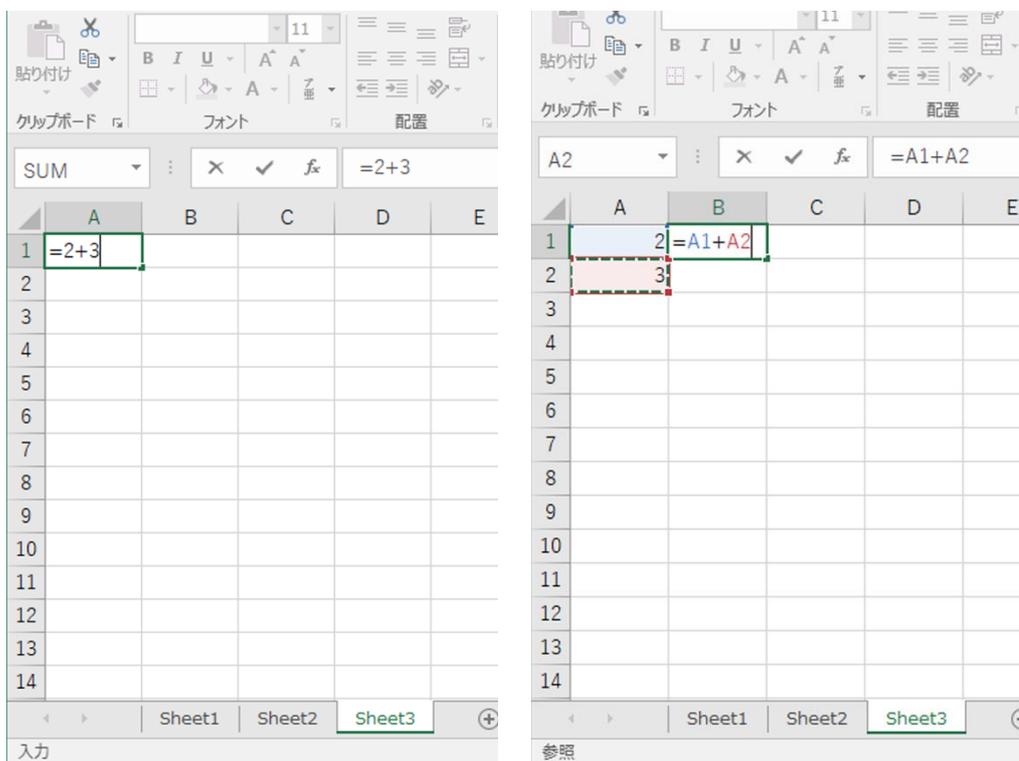
一般的にエクセルの起動時は下記のような画面になる。

四角で選択されているひとつひとつのマスを「セル」と呼ぶ。セルは番地指定ができ、左図の例では、A1のセルが指定されていることになる。

また、ウインドウ下部に「Sheet1」「Sheet2」「Sheet3」と表示されているが、これをシートと呼び、クリックすることでそれぞれのシートを選択できる。



エクセルでは、セルに数値を入力するだけでなく、計算式を入力することができる。その場合、半角で「=」から入力をはじめ、「=」の後ろに計算式を入力する。計算に用いる数値は、直接入力も可能であるし、セル番地での指定も可能である。つまり下記の2枚の画像では、いずれも2+3を計算し、結果として5がセルに表示される。



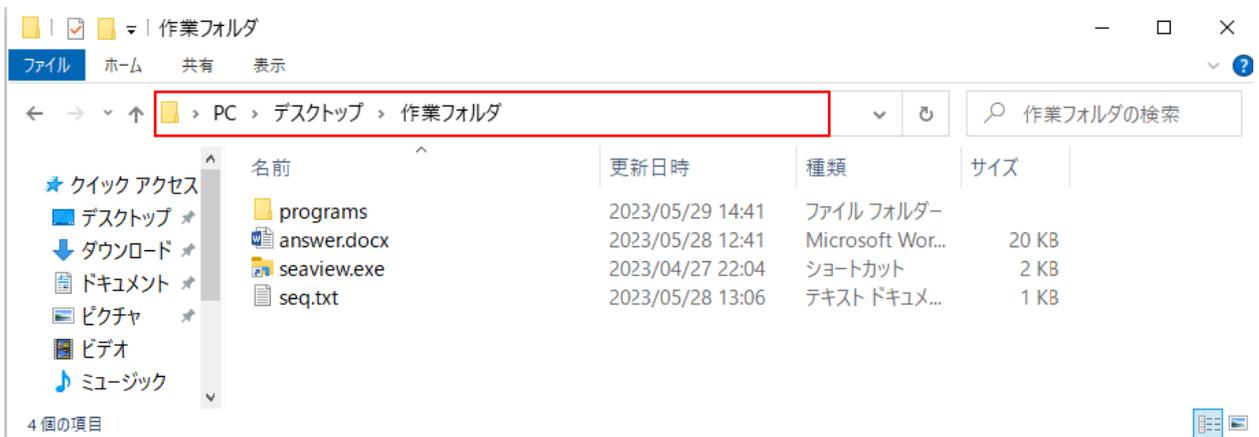
試験開始後、すぐに「ForQuestion.xlsx」をダブルクリックし、ファイルが開けることを確認すること。

解析2の準備

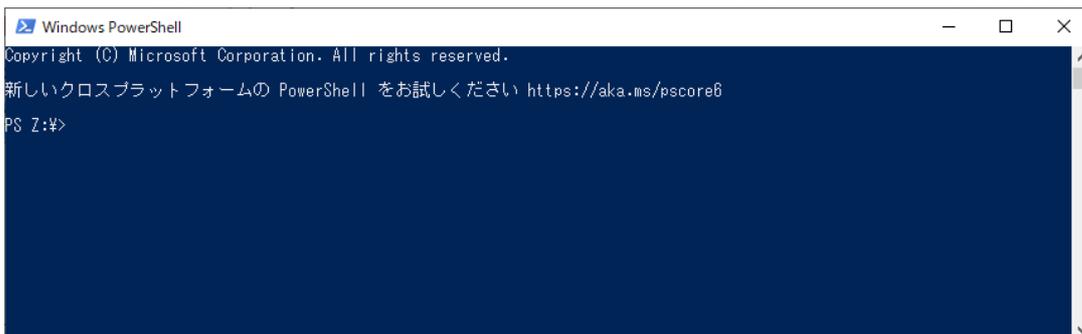
デスクトップ上に用意されている「作業フォルダ」の中に以下のファイルおよびフォルダがあることを確認する。

- programs (フォルダ)
- seaview (解析ソフト起動用ショートカット)
- seq.txt (タンパク質の配列ファイル)
- answer.docx (解答ファイル)

1. デスクトップから作業フォルダを開く。アドレスバーを右クリックし、「アドレスをコピー」をクリックする。



2. Windows キーを押してから、pw を入力し、エンターキーを押すと Windows Power Shell が起動する。

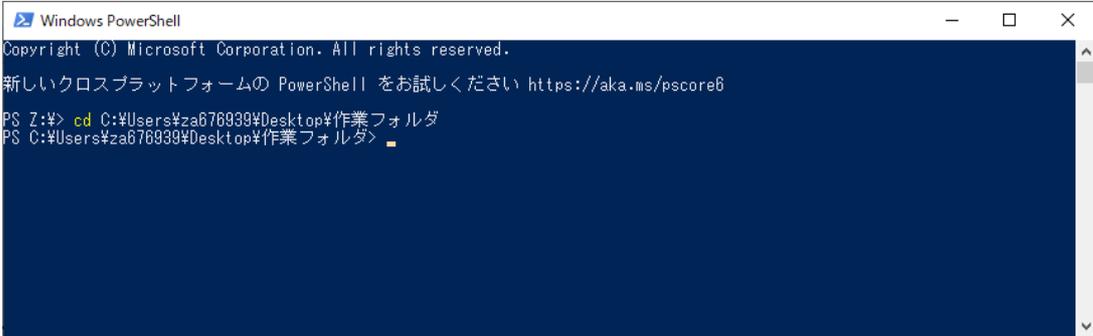


3. “cd “を入力し(cd の後ろは半角スペース)、ウインドウの上で右クリックする。



```
Windows PowerShell
Copyright (C) Microsoft Corporation. All rights reserved.
新しいクロスプラットフォームの PowerShell をお試しください https://aka.ms/pscore6
PS Z:\> cd C:\Users\za676939\Desktop\作業フォルダ
```

4. エンターキーを押す。



```
Windows PowerShell
Copyright (C) Microsoft Corporation. All rights reserved.
新しいクロスプラットフォームの PowerShell をお試しください https://aka.ms/pscore6
PS Z:\> cd C:\Users\za676939\Desktop\作業フォルダ
PS C:\Users\za676939\Desktop\作業フォルダ>
```

5. 準備完了。すぐに使わないので、最小化しておく。

ここから先は試験開始の合図があるまで
めくってはいけません

「解答始め」の合図の後、すべての解答
用紙に受験番号と名前を記入して下さい

解析 1 『PCR 増幅解析』

はじめに

近年、遺伝子発現量の解析やウイルスの感染検査等にポリメラーゼ連鎖反応(PCR)が多用されるようになってきた。PCRは耐熱性のDNAポリメラーゼ、プライマー、デオキシリボヌクレオシド三リン酸、鋳型DNAを混合し、熱変性、アニーリング(プライマー付加)、DNA合成をそれぞれ約95°C, 60°C, 72°Cで行うことを1サイクルとし、原理的には1サイクルでDNAが2倍に増幅する手法である。その一般式は下記の通り表される。

<PCRの一般式>

$$[\text{DNA}] = [\text{DNA}_0] \times (1 + e)^{\text{Ct}}$$

[DNA]: Ct サイクルの PCR 後に得られる DNA 量

[DNA₀]: 初期鋳型 DNA 量

e: 用いたプライマーの増幅効率、 $0 < e < 1$

Ct: PCR のサイクル数

この一般式をCtについて解くと以下ようになる。

$$\text{Ct} = -\frac{1}{\log(1 + e)} \times \log[\text{DNA}_0] + \frac{\log[\text{DNA}]}{\log(1 + e)} \dots \dots \textcircled{1}$$

ここでlogは10を底とした対数値を表しており、例えば、10の3乗が1000であることに対して、次のような関係が成り立つ。

$$10^3 = 1000$$

$$\log(1000) = 3$$

なお、プライマーの増幅効率eは、用いるプライマーによって異なり、理想状態ではe = 1である(サイクルごとにDNAが2倍に増幅する)。

以上を踏まえた上で、次ページからの問題に取り組みなさい。

問 1 - 1

初期鋳型DNA量を様々な濃度で用意し、ある一定のDNA量に増幅するまでに要したPCRのサイクル数をCtとしたところ、遺伝子Aと遺伝子Bについてエクセルファイルの「データ」シートのような実験結果が得られた。初期DNA量が10倍ずつプロットされているため、このグラフは見にくい。横軸の値を初期DNA量の対数値 ($\log[\text{DNA}_0]$) とするとグラフは見やすくなる。エクセルファイルの「log DNA0」シートを開き、黄色で塗りつぶしたセルに、 $\log[\text{DNA}_0]$ を計算して解答用紙に記入しなさい。なお、エクセルでは $\log(\text{値もしくはセル番地})$ とすることで対数値を計算することができる。

問 1 - 2

「log DNA0」シートで計算した値を利用して、配布したグラフ用紙に横軸 $\log[\text{DNA}_0]$ 、縦軸に遺伝子AおよびBのCt値をとったプロットを作成し、遺伝子AおよびBのグラフの傾きを求め、解答用紙に記入しなさい。

問 1 - 3

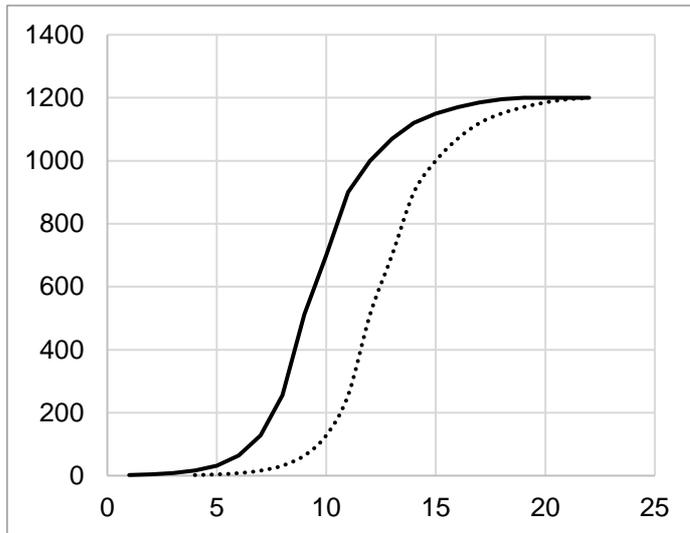
問 1 - 2 で作成したグラフは横軸 $\log[\text{DNA}_0]$ 、縦軸 Ct 値を表している。そこであらためて、説明文中のPCR一般式をCtについて解いた式①を見てみると、この式は横軸 $\log[\text{DNA}_0]$ 、縦軸 Ct 値の一次関数であることが分かる。この式と問 1 - 2 で求めた傾きをもとに、エクセルファイルの「増幅効率」シートを開き、黄色で塗りつぶしたセルに適切な値、もしくは式を入力し、遺伝子 A および B の増幅効率を計算し、解答用紙に記入しなさい。四則演算や対数の求め方などエクセルの使用方法はファイル内のシートに記載してあるので参考にすること。

問 1 - 4

遺伝子Aと遺伝子Bでは増幅効率が異なる。遺伝子AとBの初期鋳型DNA量がいずれも1分子であると仮定し、PCRを20サイクル行ったとき、増幅産物がどれだけ異なるか、その比 ($A \div B$) をエクセルファイルの「発現量の差」シートを開き、黄色で塗りつぶしたセルに計算し、解答用紙に記入しなさい。

問 1 - 5

遺伝子発現量解析に用いられる逆転写PCRは、サンプル中のRNAをDNAに逆転写した後、それを鋳型としてPCRを行うことで、サンプル中のRNA量を見積もることができる。従来の逆転写PCRではPCR後に電気泳動を行い、バンドの濃さでサンプル中のRNA量を見積もっていた。それと比べて、近年の定量逆転写PCRは、DNAに取り込まれる蛍光物質を利用しPCRのDNA増幅過程をリアルタイムに観察する



ため、より正確なサンプル中のRNA量を知ることができる。異なる2つのサンプル中に含まれるある遺伝子のRNA量を、定量逆転写PCRで解析したところ図のような結果が得られた。従来の逆転写PCRよりも定量逆転写PCRの方がDNA量を正確に見積もることができる理由について、図を参照しながら答えなさい。縦軸は増幅されたDNA量、横軸はPCRのサイクル数を示している。グラフの実線、破線は異なる2つのサンプルにおける増幅を示して

いる。

問 1 - 6

逆転写PCRでは、最初に調製したRNAにゲノムDNAが混入した場合、ゲノムDNAを鋳型としたPCRの増幅が起こることで逆転写された標的遺伝子の増幅が阻害され、その定量が不正確になる場合がある。ゲノムDNAを鋳型とした増幅は、プライマーを遺伝子上のどの位置に設計するか工夫することで、ある程度防ぐことができる。エクセルファイルの「プライマー」シートに、ある遺伝子の構造を模式的に示した。矢印は、プライマーが付加する位置と、DNA合成の向きを示している。理論的に増幅する可能性があるプライマーセットを全て挙げ、その中で定量逆転写PCRの際、ゲノムDNAからの増幅を防ぐのに最も適切なプライマーセットを答えなさい。なお、定量逆転写PCRで用いるプライマーセットで増幅する産物は、なるべく短いものが適しているものとする。

解析 2 『アミノ酸配列比較と系統樹』

問 2 - 1

配布フォルダ中の sequences.txt というファイルに各種生物 (A, B, C, D, E, F) のグロビン α サブユニットのアミノ酸配列が記載されている。以下の手順で Clustal Omega を用いて、それぞれの配列を整列(align)しなさい。以下のコマンドを Windows Power Shell で入力し、エンターを押すと実行される。ただし、 は半角スペースを意味する。このコマンドは解答ファイルにも記載されており、それを利用してよい。

```
.¥programs¥clustalo -i seq.txt -o align.txt --force --outfmt=clu
```

結果は作業フォルダ中に align.txt という名前のファイルとして出力される。この整列済みのアミノ酸配列データを開き、コピーして解答ファイルの所定の位置に張り付けなさい。

問 2 - 2

align.txt の中に示される記号の意味は以下のとおりである。

*その座位においてすべての生物種でアミノ酸が一致

:その座位においてアミノ酸が一致しないものの互いに物理化学的な性質がよく似ている

. その座位においてアミノ酸が一致しないものの互いに物理化学的な性質がある程度似ている

整列後の配列データの中で、最も高度に保存された 16 アミノ酸残基で構成される領域を黄色のハイライトで示しなさい。なお、alt キーを押しながら複数行にわたりドラッグすることにより、領域をブロック形に指定できる。

問 2 - 3

整列したアラインメントデータ(align.txt)の中で、保存性が高い(互いに類似性が高い)領域について、どのようなことが考えられるか、以下の①～④から最も適切なものを一つ選びなさい。

- ① 機能的に重要である
- ② 進化する速度が速い
- ③ 配列の信頼性が高い(正しく配列決定されている)
- ④ 機能が分かっている

問 2 - 4

本問では整列されたアミノ酸配列データ(align.txt)を用いて、系統樹構築を行う。まず seaview を起動し、「File」から「Open」をクリックし、align.txt を開きなさい。そして、「Trees」から「Distance methods」を選択し、「Go」をクリックして系統樹を作成しなさい。このとき、オプションは初期設定のままとする。「Edit」から「copy」をクリックすることで推定した系統樹の画像をコピーし、解答ファイルの所定の位置に貼り付けなさい。

問 2 - 5

A から F は実在の生物である。F はヒトであるが、その他の生物種 (A, B, C, D, E) が以下のどの生物に該当するか、以下の①～⑤の中から、答えなさい。なお、D は外群生物である。

- ① イヌ、② アフリカツメガエル、③ ゼブラフィッシュ、④ マウス、⑤ オポッサム

問 2 - 6

系統樹内における横方向に伸びた枝の長さ(横枝長とする)は両端間の変異の量を表す。しかし、ある共通祖先からこの共通祖先に由来する複数の現生種に至る横枝長が大きく異なる場合がある。つまり、分岐してから同じ時間が経っているにもかかわらず、変異の量が異なることを意味する。この場合、どのような理由が考えられるか、答えなさい。

問 2 - 7

B の祖先生物のタンパク質をコードする遺伝子が C の祖先生物に水平伝播した場合、このタンパク質の系統樹はどのような樹形になるか、答えなさい。解答は問 2 - 4 の解答である系統樹を問 2 - 7 の解答欄に手書きで写し、その系統樹に水平伝播した配列の枝を黒い点線で追加することで示しなさい。この黒い点線の先端には C' と記入しなさい。ただし、遺伝子水平伝播にかかわらず、遺伝子の欠失は起こらないものと仮定する。また、枝の長さについては問わないものとする。

問 2 - 8

問 2 - 7 のように遺伝子の水平伝播が起こり、さらに片方の遺伝子の欠失が起こった場合、その遺伝子がコードするタンパク質のアミノ酸配列で系統樹を構築すると、生物種の真の系統樹と異なった樹形になることがある。したがって、可能であれば遺伝子水平伝播が起こっていない遺伝子がコードしているタンパク質を用いて種の系統樹を推定することが望ましい。しかし、実際には対象とする遺伝子の水平伝播の有無が分からない場合も多いため、信頼性の高い種の系統樹を推定するには、工夫が必要になる。どのような工夫が望ましいか、理由とともに答えなさい。

これで問題は終了です