





# 業 務 仕 様 書

1. 件 名 次世代シーケンス解析支援および開発支援業務
2. 業務目的 本業務は、次世代シーケンサーを用いたメタゲノム解析に係る塩基配列データを効率よく解析するため、バイオインフォマティクス技術を用いて、解析とその検証を繰り返しながら解析技術の開発を行う。
3. 業務場所 神奈川県横浜市金沢区福浦 2-12-4  
国立研究開発法人水産研究・教育機構  
横浜庁舎及び請負業者指定場所
4. 業務期間 自) 令和3年6月 1日  
至) 令和4年1月31日

5. 業務内容 **【業務概要】**

海水中に出現する動植物プランクトン・細菌・生物群の分布範囲、季節性、出現頻度等の海域特性をゲノム情報から明らかにすることを目的として、メタゲノム解析技術により、試料中の微生物のDNAを混合物として抽出し、このDNA集合体の塩基は塩基配列を解読することで、試料中に含まれる微生物の種類やその存在比率を推定する。次世代シーケンサーから得られる膨大な塩基配列データからプログラムを用いて配列に同じ文字やパターンを見つけ、データベース上の既知の生物種と照合し種を判別するが、高度な解析の実施や、研究テーマに沿った結果を得るためには、既存ソフトウェアの機能では十分でなく、解析者がプログラミングする必要があり、プログラミングを習熟している生物学研究者、バイオインフォマティクスの専門技術を有する者による研究課題に応じたオーダーメイド解析及び解析方法の開発を行う研究支援業務である。

**【作業内容】**

- (1) 次世代シーケンサーIllumina 社の MiSeq 300PE (V3) platform を用いたプランクトン・細菌・魚卵メタゲノム解析による種同定技術の開発および生物多様性比較解析のためのプラットフォーム構築に関する

業務のうち、次世代シーケンスデータの配列エラーの検出やノイズ除去法の開発に関する業務

- (2) 同業務のうち、次世代シーケンスデータを用いたクラスタリングや blast 検索による種同定技術の開発に関する業務
- (3) 同業務のうち、種同定結果のデータ集計や表およびグラフ作成に関する業務
- (4) 植物プランクトン（珪藻）の全ゲノム解析業務のうち、次世代シーケンスデータを用いた Newbler 等のソフトを用いたアセンブル作業に関する業務
- (5) 同業務のうち、Augustus 等のソフトを用いた遺伝子予測解析に関する業務
- (6) 上記（1）から（5）に付帯する業務
- (7) 電話やメールまたはリモートアクセスにより、担当職員からの質問に対応すること。
- (8) 解析結果の報告は、エクセル、ワード等電子ファイル形式でメールにより担当職員へ報告すること。
- (9) 各作業の詳細な内容は、適宜、担当職員と協議し決定するものとする。

#### 【作業従事者に求める技術等】

本業務は生物学のデータを情報科学の手法によって解析するものであり、その実施には、生物学と情報科学の知識が必要となり、次世代シーケンサーを利用したゲノム解析では、大きな計算能力を要求される課題が多く存在するため、作業従事者に以下の技術、経験を備えていることを求める。

- (1) ゲノム解析及びメタゲノム解析の基本的な流れ、ゲノムに関する基本的な用語、及び、解析に関する技術用語を理解していること。
- (2) 次世代シーケンサー由来のデータを扱った業務に従事した経験があること。
- (3) ゲノム解析に関する実務的なソフトウェアの設計または開発経験があること。
- (4) Linux もしくは Unix の OS の使用経験があり、C, Java, perl, ruby, python 等のコンピュータ言語について専門的な知識があること。
- (5) 以下の何れかのバイオインフォマティクスソフトウェアを利用したデータ解析手法の開発経験があること。

(バイオインフォマティクスソフトウェア)

ABySS, ALLPATH-LG, Augustus, bbtools, bcftools, BEDtools, BESST, Bioconductor, BioJava4, BioPerl, Biopython, BioRuby, blasr, BLAT, busco, BWA, canu, cegma, Clustal-omega, Cytoscape, Edena, EMBOSS, Exonerate, Falcon-Unzip, FastQC, FASTX-Toolkit, FreeBayes, GATK4, GMclosure, IGV, jabba, karect, MAFFT, minimap2, mira4, MUMmer, ncbiutils, Perl, Picard, platanus, plink, Python, RepeatMasker, RepeatScout, RepeatModeler, Ruby, SAMtools, SMRT Link (Falcon, HGAP), SnpEff, SOAP, SOAP denovo, SPAdes, SRA Toolkit, TRF, supernova2, Trimmomatic, Trinity, VarScan, vcftools, Velvet, vsearch

#### 【実施体制等】

- (1) 本業務を実施するための人員および組織体制を選任、確保し、契約期間にわたりその組織体制を維持すること。
- (2) 請負者が選任した者は、発注者の依頼に基づき毎週3日程度（業務期間中合計100日間）業務に従事し、そのうち1日間は訪問のうえ業務を行うこと。また、これとは別に毎月1日程度（業務期間中合計8日間）の訪問のうえ、担当職員と結果解釈についてのディスカッションを行い、請負者の保有する施設及び当研究所にて次世代シーケンサーを用いたデータ解析及び開発支援等を行い、結果を報告すること。
- (3) 業務従事日及び訪問日は、担当職員と協議し決定するものとする。
- (4) 請負者は業務の実施に当たって毎月20日までに翌月分の作業実施計画表を提出し承認を得ること。
- (5) 当研究所が指定する休日（土曜日、日曜日、国民の祝日、12月29日から翌年の1月3日までの日）は訪問作業実施日から除く。

#### 【設備等】 データ解析サーバー

ナベインターナショナル社製

TS-R2695v3A (2) :Takeru For Sequencer V

CPU : Intel 14-core Xeon e5-2695 v3 ×2

Memory : 128GB (16GB×8) DDR4-2133 ECC Registered

OS : CentOS7 for x86\_64

6. 作業の確認 請負者は月々の業務が完了後、速やかに「業務完了報告書」を提出し検査を受けること。

7. 特記事項
- (1) 請負者は、業務に関連する全てのデータについて、発注者の許可なく当該業務の処理以外の目的に使用し、又は外部に提供してはならない。
  - (2) 請負者は、請け負った一連の業務が完了した時点で、担当職員から借用したデータは廃棄するものとする。
  - (3) 請負者は、業務上知り得た機密事項を一切他に漏洩してはならない。
8. その他
- 詳細については担当職員の指示に従うものとする。