






NOTE

**Cahier des Charges
N° IBFJ-2022-002**

PRESTATIONS DE BIOINFORMATIQUE ET SEQUENCAGE

	NOM	FONCTION	VISA	DATE
REDACTION	MENU Elisabeth/BEIGN ON Anne-Sophie	Chercheurs		23/06/2025
REDACTION	MANGEOT Isabelle KRZACZKOWSKI Lucie	Chargée de programmes / Responsable de programmes partenariaux-valo		23/06/2025
VALIDATION	LE GRAND Roger	Directeur IDMIT		

OBJET

Le présent Cahier des Spécifications Techniques Particulières (CSTP) a pour objet de détailler les prestations d'analyse bioinformatique et de séquençage que le Commissariat à l'Energie Atomique et aux énergies alternatives – Centre de Paris-Saclay souhaite confier au Titulaire pour l'Institut de Biologie François Jacob.

PRESENTATION DE LA PRESTATION

La Prestation demandée consiste en la réalisation par le Titulaire de séquençages par différentes technologies, de l'analyse de données bioinformatiques en lien avec différentes technologies omiques et de calcul haute performance (HPC) pour l'Institut de Biologie François Jacob (Centre CEA de Paris-Saclay – Site de Fontenay-aux-Roses (FAR)).

a. Locaux mis à disposition

Les locaux dévolus au Titulaire sont situés dans le Département IDMIT du centre CEA de Fontenay-aux-Roses (18 route du Panorama, 92260 Fontenay-aux-Roses) et comprennent :

- 1 bureau au bâtiment 62
- 1 zone de laboratoire L1

Le Titulaire aura également accès au sein du Département IDMIT aux laboratoires L1, L2 et L3 situés dans le bâtiment (suivant les besoins des prestations et autorisations délivrées par l'ingénieur sécurité d'IDMIT).

Les locaux auxquels le Titulaire a accès sont pré-équipés en connectivité réseau.

L'accès aux locaux est protégé. Les personnels dédiés à la réalisation de la Prestation se verront attribuer un badge d'accès nominatif par le CEA, sous réserve de recevoir avant l'arrivée des nouveaux personnels, les documents demandés par le CEA pour l'autorisation d'entrée sur site.

b. Equipements et consommables

Le CEA met à disposition du Titulaire les équipements situés dans le bâtiment IDMIT à l'exclusion des équipements dédiés aux activités de zootechnie. Notamment, en ce qui concerne la Prestation :

- 4150 TapeStation, Agilent
- TECAN, Spark
- Chromium Controller, 10X
- Réfrigérateur 4°C (partagé avec les équipes du CEA)
- Congélateur -20°C (dédié au Titulaire)
- Congélateur -80°C (partagé avec les équipes du CEA)

Le Titulaire devra installer son/ses séquenceur(s) dans le bâtiment IDMIT. Pour toutes les tâches des prestations qui seront réalisées à l'extérieur du bâtiment IDMIT, le transport des échantillons sera à la charge du Titulaire. Les conditions de transport des échantillons seront précisées pour chaque prestation.

Une partie des prestations étant possiblement réalisée sur des équipements du Département IDMIT du CEA, le Titulaire devra justifier de ses compétences pour l'utilisation de ces équipements et les analyses demandées. La maintenance et l'entretien de ces équipements sont à la charge du CEA.

Le CEA fournit également les consommables « de base », non spécifiques à la réalisation de la Prestation. Par exemple (non exhaustif) : cônes, tubes, produits chimiques, EPI.

Le Titulaire garantit disposer ou accéder aux équipements nécessaires à la réalisation de la Prestation : séquenceurs, petits matériels de paillasse (dont pipettes, centrifugeuse, thermocycleurs, et matériel de dosage fluorimétrique) et consommables spécifiques. Le Titulaire prendra en charge les maintenances et qualifications de ces appareils le cas échéant.

c. Gestion des reliquats d'échantillons et des déchets

A l'issue des prestations de 1 à 5, le Titulaire informera le CEA des échantillons restants qui décidera de leur destruction ou de leur récupération.

Le CEA se chargera de la destruction des échantillons et fluides usagés générés par les prestations.

d. Réseau informatique et logiciels

Les personnels du Titulaire disposent d'une identité numérique individuelle permettant une connexion sécurisée aux postes informatiques et logiciels du CEA.

Le Département IDMIT du CEA dispose du logiciel BATLab développé sur-mesure, géré et maintenu par une équipe dédiée. Ce logiciel est accessible, via un compte personnel et protégé, à l'ensemble du personnel du Département IDMIT. Il consiste en un système de gestion de données de laboratoire (LIMS) utilisé notamment pour la gestion et le traitement des données scientifiques produites. Par ailleurs, dans le cadre de la démarche qualité du Département IDMIT, BATLab répertorie les procédures, modes opératoires et autres documents consultables par tous sur lesquels s'appuie chaque processus. Les personnels mis à disposition par le Titulaire pour réaliser la Prestation au sein du bâtiment IDMIT devront utiliser ce logiciel. Le CEA apportera une formation et un accompagnement à l'utilisation du logiciel BATLab pendant une semaine. A l'issue de cette période, le CEA accordera une qualification nominative avec un PV.

Par ailleurs, pour les prestations 1 à 5 ci-après, il est prévu un transfert des données brutes de séquençage vers les serveurs du CEA. Ce transfert sera à la charge du Titulaire.

e. Ressources humaines

Le Titulaire s'engage à affecter à l'exécution des activités de la Prestation, sur le site de Fontenay-aux-Roses, les personnels mentionnés dans son mémoire technique. Il garantit les compétences, l'expérience et les qualifications professionnelles de ces intervenants. Notamment, les compétences doivent être adaptées aux modèles étudiés : humain, primate non humain, murin.

Le CEA devra être informé de toute modification de l'équipe.

f. Qualité

Le Département IDMIT du CEA est engagé dans une démarche qualité et certifié ISO 9001 depuis 2017. Le Titulaire est tenu de respecter le système de management de la qualité mis en place au sein du Département. La certification ISO 9001 du Titulaire est requise.

g. Activités de la Prestation

La Prestation comprend les activités suivantes :

1) Prestations de séquençage de génome entier (sur ADN)

Le Titulaire devra réaliser des analyses de séquençage d'ADN sur des échantillons biologiques (microorganismes, cellules, liquides biologiques, tissus, broyats d'organes) infectés ou non par des pathogènes de classe 2 ou de classe 3 ou à partir du matériel biologique extrait de ces types d'échantillons par le CEA.

Le CEA pourra fournir jusqu'à 100 échantillons sur la durée de la Prestation.

Dès la première prestation, le Titulaire devra nommer et affecter un pilote opérationnel en charge des prestations « séquençage de génome entier (sur ADN) ». Ce dernier veillera notamment au respect des délais et des exigences de qualité. Il est attendu que le Titulaire apporte une aide dans la conception des projets notamment en conseillant sur le design expérimental le plus approprié pour répondre à la question biologique. Il assurera un accompagnement tout au long de la prestation « séquençage de génome entier » avec des réunions régulières d'échanges et de discussions autour des résultats (même intermédiaires).

La prestation de séquençage de génome entier (sur ADN) comprend :

- Extraction d'ADN fragments à partir de différents échantillons, le cas échéant
- Quantification des ADN extraits par dosage fluorimétrique
- Contrôle qualité des ADN extraits par électrophorèse automatisée
- Séquençage des ADN extraits par technologie Nanopore sur flowcell MinION ou PromethION, avec et sans multiplexage. La librairie Nanopore de séquençage génome entier utilisée devra être adaptée aux échantillons (petits organismes ou grands organismes)
- Transfert par sftp et stockage sécurisé des données brutes (fastq et rapport de séquençage)
- Analyses bioinformatiques à façon :
 - Niveau 1 (inclus dans cette prestation) : Contrôle qualité des données

Pour cette prestation, le Titulaire proposera deux niveaux d'analyses supplémentaires :

- Analyses intermédiaires (niveau 2) :
 - Assemblage
 - Comparaison d'échantillons multiples et avec des génomes de référence, notamment recherche de mutations par rapport à des génomes de référence
- Analyses poussées (niveau 3) :
 - Annotations
 - Analyses multiparamétriques intégrant les Omics et d'autres variables d'intérêt

Pour chaque projet, en fonction des besoins associés exprimés par le CEA le Titulaire évaluera le nombre de jours nécessaires pour réaliser les analyses approfondies souhaitées, sur la base du tarif journalier indiqué dans le BPU.

- Rendu des livrables :
 - Analyses de niveau 1 : Rapport de qualité de séquençage, Fichiers bruts de séquençage (FASTQ), Procédures et modes opératoires des activités réalisées par le Titulaire, Transmission de l'ensemble des données brutes dans les 7 jours ouvrés à partir de la fin du séquençage
 - Analyses de niveau 2 : Rendu de résultats sous forme de rapport avec des graphiques utilisables pour des publications (le cas échéant) au plus tard 1 mois après la génération de l'ensemble des données de séquençage du projet
 - Analyses de niveau 3 : Rendu de résultats sous forme de rapport avec des graphiques utilisables pour des publications (le cas échéant) dans les 2 mois après génération de l'ensemble des données de séquençage du projet
- Gestion, sauvegarde et archivage :
 - des données brutes pendant 1 an après séquençage
 - des données analysées pendant 1 an après le rendu des livrables

2) Prestations de transcriptomique (bulk)

Le Titulaire devra réaliser des analyses de transcriptomique sur des échantillons biologiques (cellules, liquides biologiques, tissus, broyats d'organes), les échantillons étant frais, congelés ou fixés en paraffine. Le Titulaire pourra également être amené à réaliser les analyses à partir du matériel biologique extrait par le CEA.

Le CEA pourra fournir jusqu'à 400 échantillons sur la durée de la Prestation. Le Titulaire doit être capable d'analyser des quantités d'ARN inférieures à celles indiquées par le protocole fournisseur. Il devra également adapter le cas échéant les solutions (les protocoles et les outils bioinformatiques) aux modèles étudiés au CEA (murin, simien, humain).

Dès la première prestation, le Titulaire devra nommer et affecter un pilote opérationnel en charge des prestations « transcriptomiques ». Ce dernier veillera notamment au respect des délais et des exigences de qualité. Il est attendu que le Titulaire apporte une aide dans la conception des projets notamment en conseillant sur le design expérimental le plus approprié pour répondre à la question biologique. Il assurera un accompagnement tout au long de la prestation « transcriptomique » avec des réunions régulières d'échanges et de discussions autour des résultats (même intermédiaires).

La prestation de transcriptomique (bulk) comprend :

- Extraction ARN, le cas échéant, à partir de différents échantillons biologiques dont de faibles volumes de sang (entre 50 et 200 µL) avec optimisation du protocole.
- Quantification des ARN extraits par dosage fluorimétrique
- Contrôle qualité des ARN extraits par électrophorèse automatisée
- Séquençage
 - Pré-requis : ARN avec RIN >7 et quantité adaptée au protocole
 - Par technologie Nanopore, avec multiplexage, sur flowcell MinION ou PromethION. Un minimum de 600 000 reads par échantillon est requis pour l'analyse.
 - Par technologie Illumina, avec multiplexage. Sur NextSeq 2000, d'après les recommandations Illumina, 50M de reads par échantillon sont nécessaires pour le protocole Total RNAseq et 25M pour le protocole mRNA et ≥85% de reads avec un score qualité > 30
- Transfert par sftp et stockage sécurisé des données brutes (fastq et rapport de séquençage)

- Analyses bioinformatiques à façon
 - Niveau 1 (inclus dans cette prestation) :
 - Contrôle qualité des données
 - Critère qualité du niveau 1 : > 70% de reads alignés
 - Exploration des données (ACP/Heatmaps)

Pour cette prestation, le Titulaire proposera deux niveaux d'analyses supplémentaires :

- Analyses intermédiaires (niveau 2) : analyse différentielle (comparaison d'échantillons) comprenant :
 - Correction de l'effet batch si nécessaire
 - Comparaison d'échantillons avec recherche des gènes différentiellement exprimés
 - Génération des graphiques
- Analyses poussées (niveau 3) incluant
 - l'enrichissement fonctionnel (GSEA) et la recherche de gènes spécifiques dans les sous-populations d'intérêt.
 - Analyses multiparamétriques intégrant les Omics et d'autres variables d'intérêt

Pour chaque projet, en fonction des besoins associés le Titulaire évaluera le nombre de jours nécessaires pour réaliser les analyses approfondies souhaitées, sur la base du tarif journalier indiqué dans le BPU.

- Rendu des livrables
 - Niveau 1 :
 - Rapport de qualité de séquençage incluant les séquences non alignées
 - Fichiers bruts de séquençage (FASTQ)
 - Procédures et modes opératoires des activités réalisées par le Titulaire
 - Transmission de l'ensemble des données brutes dans les 7 jours ouvrés à partir de la fin du séquençage
 - Niveau 2 : dans les 2 mois après génération de l'ensemble des données de séquençage du projet
 - Rendu de résultat sous forme de rapport avec des graphiques utilisables pour des publications (le cas échéant)
 - Rendu de résultats sous forme de gènes différentiellement exprimés entre les conditions comparées (format Excel)
 - Niveau 3 : dans les 3 mois après génération de l'ensemble des données de séquençage du projet
 - Rendu de résultat sous forme de rapport avec des graphiques utilisables pour des publications (le cas échéant)
 - Rendu de résultats sous forme de gènes différentiellement exprimés entre les conditions comparées (format Excel)
- Gestion, sauvegarde et archivage :
 - des données brutes pendant 1 an après séquençage
 - des données analysées pendant 1 an après le rendu des livrables

3) Prestation de séquençage Single cell RNAseq

Le Titulaire réalisera des analyses de séquençage single cell RNAseq à partir d'échantillons biologiques, en utilisant les équipements du CEA dans des environnements de sécurité de niveau BSL2 ou BSL3.

La maîtrise des technologies 10x Genomics et des outils bioinformatiques associés est indispensable.

Le CEA pourra fournir jusqu'à 50 échantillons sur la durée de la Prestation.

Dès la première prestation, le Titulaire devra nommer et affecter un pilote opérationnel en charge des prestations « séquençage Single cell RNAseq ». Ce dernier veillera notamment au respect des délais et des exigences de qualité. Il est attendu que le Titulaire apporte une aide dans la conception des projets notamment en conseillant sur le design expérimental le plus approprié pour répondre à la question biologique. Il assurera un accompagnement tout au long de la prestation « séquençage Single cell RNAseq » avec des réunions régulières d'échanges et de discussions autour des résultats (même intermédiaires).

Le séquençage Single cell RNAseq inclut la génération de librairie 10x Genomics, réalisée par le CEA ou par le Titulaire.

La prestation de séquençage Single cell RNAseq comprend :

- Contrôle qualité par électrophorèse automatisée
- Séquençage à haut débit sur NextSeq2000 avec multiplexage. Un minimum 20 000 reads/cellules est nécessaire pour l'analyse avec au moins 85% des bases avec une qualité supérieure à Q30. La possibilité de réaliser un séquençage shallow serait un plus et fera l'objet d'un BPU spécifique.
- Transfert par sftp et stockage sécurisé des données brutes (fastq et rapport de séquençage)
- Analyses bioinformatiques à façon :
 - Analyses primaires (niveau 1) : utilisation de Cell Ranger sur serveur
 - Critère qualité du niveau 1 suivant les recommandations 10X : « Reads mapped confidently to transcriptome » >30%

Pour cette prestation, le Titulaire proposera deux niveaux d'analyses supplémentaires :

- Analyses intermédiaires (niveau 2) : clustering des populations par groupe et identification des marqueurs différenciants pour chaque cluster en utilisant des pipelines bioinformatiques adaptés à façon
- Analyses poussées (niveau 3) incluant
 - la recherche de gènes spécifiques dans les sous-populations d'intérêt utilisant des pipelines bioinformatiques adaptés à façon
 - assignation des populations selon une liste de gènes d'intérêt
 - inférence de trajectoire

Pour chaque projet, en fonction des besoins associés le Titulaire évaluera le nombre de jours nécessaires pour réaliser les analyses approfondies souhaitées, sur la base du tarif journalier indiqué dans le BPU.

- Rendu des livrables
 - Analyses de niveau 1 : Rapport de qualité de séquençage, Fichiers bruts de séquençage (FASTQ), Procédures et modes opératoires des activités réalisées par

le Titulaire, Transmission de l'ensemble des données brutes dans les 7 jours ouvrés à partir de la fin du séquençage

- Analyses de niveau 2 : Rendu de résultats sous forme de gènes différentiellement exprimés par clusters (format Excel), Rendu de résultats sous forme de rapport avec des graphiques utilisables pour des publications (le cas échéant) dans les 2 mois après génération de l'ensemble des données de séquençage du projet
 - Analyses de niveau 3 : Rendu de résultats sous forme de gènes différentiellement exprimés par sous-populations cellulaires (format Excel), Rendu de résultats sous forme de rapport avec des graphiques utilisables pour des publications (le cas échéant) dans les 3 mois après génération de l'ensemble des données de séquençage du projet
- Gestion, sauvegarde et archivage :
 - des données brutes pendant 1 an après séquençage
 - des données analysées pendant 1 an après le rendu des livrables

4) Prestations d'analyses du Microbiote :

Le Titulaire réalisera des analyses de microbiote à partir d'échantillons biologiques infectés ou non (fèces, BAL, fluides vaginaux, swabs) ou à partir du matériel biologique extrait de ces types d'échantillons par le CEA, en utilisant des technologies de séquençage Illumina ou Nanopore. Le Titulaire pourra analyser jusqu'à 1 000 échantillons sur la durée de la Prestation. Le Titulaire doit être capable d'analyser des faibles quantités d'ADN (entre 1 et 12,5 ng) sur le protocole Illumina V3-V4.

Dès la première prestation, le Titulaire devra nommer et affecter un pilote opérationnel en charge des prestations « analyses du Microbiote ». Ce dernier veillera notamment au respect des délais et des exigences de qualité. Il est attendu que le Titulaire apporte une aide dans la conception des projets notamment en conseillant sur le design expérimental le plus approprié pour répondre à la question biologique. Il assurera un accompagnement tout au long de la prestation « analyses du Microbiote » avec des réunions régulières d'échanges et de discussions autour des résultats (même intermédiaires).

La prestation d'analyses du Microbiote comprend :

- Quantification des ADN extraits par dosage fluorimétrique
- Contrôle qualité par électrophorèse automatisée
- Séquençage :
 - Par technologie Illumina :
 - metabarcoding 16S V3-V4 avec flowcell 2x300 cycles. Un minimum de 100 000 reads par échantillon est requis en sortie de séquenceur avec au moins 70% des bases avec une qualité supérieure à Q30. L'option d'analyse à partir de faibles quantités fera l'objet d'une ligne spécifique dans le BPU.
 - Shotgun avec option déplétion de l'ADN de l'hôte (majoritairement humain ou macaque) sur flowcell 2x150 cycles.
 - Pour une analyse de la taxonomie : 4 millions de reads bruts sont requis avec au moins 80% des bases avec une qualité supérieure à Q30.

- Pour une analyse des gènes de résistance aux antibiotiques : 35 millions de reads bruts sont requis avec au moins 85% des bases avec une qualité supérieure à Q30.

Par technologie Nanopore : Séquençage Full 16S avec multiplexage sur flowcell MinION ou PromethION. Un minimum de 25 000 reads par échantillon est requis pour l'analyse.

- Transfert par sftp et stockage sécurisé des données brutes (fastq et rapport de séquençage)
- Analyses bioinformatiques via une interface graphique :
 - Niveau 1 (inclus dans cette prestation) : contrôle qualité des données

Pour cette prestation, le Titulaire proposera deux niveaux d'analyses supplémentaires :

- Analyses intermédiaires (niveau 2) comprenant via une interface graphique :
 - traitement des séquences, assignation des OTU, ,
 - génération de graphiques dynamiques (visualisation à différents niveaux taxonomiques)
 - Calculs d'alpha et betadiversité,
 - Possibilité de fournir des metadonnées pour visualiser les différents groupes
 - Résultats et figures exportables aux formats xlsx et png
- Analyses bioinformatiques à façon (niveau 3) réalisées sur des serveurs sécurisés: recherche de corrélations, comparaison d'échantillons, analyses multiparamétriques intégrant les Omics et d'autres variables d'intérêt. Une analyse différentielle intégrant la comparaison de conditions pourra être souhaitée dans certains cas par le CEA.

Pour chaque projet, en fonction des besoins associés le Titulaire évaluera le nombre de jours nécessaires pour réaliser les analyses approfondies souhaitées, sur la base du tarif journalier indiqué dans le BPU.

- Rendu des livrables
 - Analyses de niveau 1 : Rapport de qualité de séquençage, Fichiers bruts de séquençage (FASTQ), Procédures et modes opératoires des activités réalisées par le Titulaire, Transmission de l'ensemble des données brutes dans les 7 jours ouvrés à partir de la fin du séquençage
 - Analyses de niveau 2 : Rendu des assignations des OTU, des résultats de calculs d'alpha et beta diversité sous format Excel, Rendu de résultats sous forme de rapport avec des graphiques utilisables pour des publications au plus tard 1 mois après la génération de l'ensemble des données de séquençage du projet
 - Analyses de niveau 3 : Rendu de résultats sous forme de rapport au format Excel et avec des graphiques utilisables pour des publications dans les 2 mois après génération de l'ensemble des données de séquençage du projet
- Gestion, sauvegarde et archivage :
 - des données brutes pendant 1 an après séquençage
 - des données analysées pendant 1 an après le rendu des livrables

5) Prestation d'analyses de séquences virales

Le Titulaire réalisera ces analyses à partir de différents échantillons biologiques (cellules, liquides biologiques, tissus, broyats d'organes, stock viral) infectés par des pathogènes de classe 2 ou de classe 3 ou à partir du matériel biologique extrait de ces types d'échantillons par le CEA.

Le Titulaire pourra analyser jusqu'à 50 échantillons sur la durée de la Prestation.

Dès la première prestation, le Titulaire devra nommer et affecter un pilote opérationnel en charge des prestations « analyses de séquences virales ». Ce dernier veillera notamment au respect des délais et des exigences de qualité. Il est attendu que le Titulaire apporte une aide dans la conception des projets notamment en conseillant sur le design expérimental le plus approprié pour répondre à la question biologique. Il assurera un accompagnement tout au long de la prestation « analyses de séquences virales » avec des réunions régulières d'échanges et de discussions autour des résultats (même intermédiaires).

La prestation d'analyses de séquences virales comprend :

- Extraction d'ADN ou d'ARN à partir de différents échantillons biologiques, le cas échéant
- Quantification des ADN/ARN extraits par dosage fluorimétrique
- Contrôle qualité des ADN/ARN extraits par électrophorèse automatisée
- Séquençage
 - A partir d'amplicons viraux et par technologie Nanopore, avec multiplexage, sur flowcell MinION.
- Analyses bioinformatiques via une interface graphique :
 - Niveau 1 (inclus dans cette prestation) : Contrôle qualité des données

Pour cette prestation, le Titulaire proposera deux niveaux d'analyses supplémentaires :

- Analyses intermédiaires (niveau 2) comprenant via une interface graphique :
 - Identification des virus présents dans l'échantillon
 - Identification des variants viraux et annotations
 - Résultats et figures exportables aux formats xls et png
- Analyses bioinformatiques à façon (niveau 3) réalisées sur des serveurs sécurisés: recherche de corrélations, comparaison d'échantillons, analyses multiparamétriques intégrant les Omics et d'autres variables d'intérêt
- Rendu des livrables :
 - Analyses de niveau 1 : Rapport de qualité de séquençage, Fichiers bruts de séquençage (FASTQ), Procédures et modes opératoires des activités réalisées par le Titulaire, Transmission de l'ensemble des données brutes dans les 7 jours ouvrés à partir de la fin du séquençage
 - Analyses de niveau 2 : Rendu des résultats d'analyse sous format Excel, Rendu de résultats sous forme de rapport avec des graphiques utilisables pour des publications au plus tard 1 mois après génération de l'ensemble des données de séquençage du projet
 - Analyses de niveau 3 : Rendu de résultats sous forme de rapport avec des graphiques utilisables pour des publications dans les 2 mois après génération de l'ensemble des données de séquençage du projet

- Gestion, sauvegarde et archivage :
 - des données brutes pendant 1 an après séquençage
 - des données analysées pendant 1 an après le rendu des livrables

6) Prestation d'analyses de données de cytométrie et omiques

Le Titulaire réalisera des analyses de données de cytométrie et omiques, en utilisant des méthodes bioinformatiques avancées. Le CEA évalue son besoin à 40 jours maximum par an. La maîtrise des langages R et Python ainsi que l'administration de systèmes (Linux) et la maintenance de serveurs sont indispensables.

Dès le démarrage de la prestation, le Titulaire devra nommer et affecter un Pilote opérationnel qui veillera notamment au respect des délais et des exigences de qualité. Il est attendu que le Titulaire apporte une aide dans la conception des projets notamment en conseillant sur le design expérimental le plus approprié pour répondre à la question biologique. Il assurera un accompagnement tout au long de la prestation « analyses de données de cytométrie et omiques » avec des réunions régulières d'échanges et de discussions autour des résultats (même intermédiaires).

La prestation d'analyses de données de cytométrie et omiques comprend :

- Réaliser une analyse bioinformatique sur les données de cytométrie de grande dimension en cytométrie de masse et en cytométrie de flux et des données Omics
- Construire des pipelines d'analyses non supervisées adaptés aux jeux de données
- Proposer/utiliser la méthode d'analyse optimale en fonction du jeu de données à l'aide d'algorithmes déjà disponibles ou développés sur-mesure
- Analyses statistiques
 - comparaisons entre groupes d'individus ou en cinétique pour l'abondance de cellules des clusters ou régions
 - comparaisons de phénotypes entre clusters ou régions pour préciser les annotations
 - analyses de corrélations
 - analyses multiparamétriques intégrant les Omics et d'autres variables d'intérêt (imagerie par exemple)
- Rendu de livrables :
 - Procédures et méthodes utilisées par le Titulaire
 - Rendu de résultats sous forme de rapport avec des graphiques utilisables pour des publications (le cas échéant)
- Gestion, sauvegarde et archivage :
 - des données brutes pendant 1 an après séquençage
 - des données analysées pendant 1 an après le rendu des livrables

7) Prestation d'accompagnement à l'analyse des données

A la demande du CEA, le Titulaire organisera des sessions de sensibilisation sur l'interprétation des résultats et l'utilisation des outils bioinformatiques associés. Il pourra également fournir une prestation de conseils et de support en bioinformatique et biostatistiques, HPC et Machine learning.

Le CEA évalue son besoin à une semaine maximum sur la durée de la Prestation.

Contraintes expérimentales et techniques

- Stockage des données en France
- Encryptions des données
- Transfert des données de séquençage vers les serveurs d'analyse du CEA
- Les analyses demandées pourront être réalisées sur des échantillons infectés avec des pathogènes de classe 2 ou de classe 3 pour le risque biologique.
- Lorsque des analyses de transcriptomique (prestation 3) peuvent être réalisées sur des échantillons frais, le Titulaire sera mobilisé selon le planning établi par le CEA (hors week-end et jours fériés).
- Les analyses de transcriptomique et sur le microbiote (prestations 2 et 4) doivent pouvoir être réalisées sur de faibles quantités de matériel. Le Titulaire doit disposer de protocoles adaptés.
- Les prestations 2 à 5 peuvent être réalisées sur plusieurs mois/années. Le Titulaire devra préciser les protocoles qualité mis en œuvre pour tenir compte de la correction entre les groupes d'échantillons.

Délais

Les délais associés à la remise des livrables sont précisés pour chaque prestation. Dans le cadre de projets de recherche, le CEA s'engage à déclencher les BPU un mois avant le démarrage souhaité de la prestation. En cas d'urgence (exemple : pandémie), compte tenu de la mission d'Infrastructure Nationale en Biologie-Santé (INBS) d'IDMIT, il sera attendu que le Titulaire soit réactif.

PROPOSITION FINANCIERE POUR LES PRESTATIONS ENVISAGEES

Pour chacune des prestations proposées (1, 2, 3, 4, 5, 6 et 7), le Titulaire fournira un bordereau de prix pour le nombre d'échantillons ou le temps d'analyse précisé. Comme indiqué dans le détail des différentes prestations et suivant les projets associés, le CEA pourra souhaiter des niveaux d'analyse plus poussés. Le Titulaire devra fournir un tarif journalier pour la réalisation de ces analyses approfondies (niveaux 2 et 3).

Par ailleurs, pour les prestations 1, 2, 4 et 5, il fournira des propositions de prix avec l'extraction en option (réalisée par le Titulaire ou par le CEA) pour des extractions d'ADN/ARN selon 2 critères :

- Protocoles « classiques »
- Protocoles « spécifiques » (type PAXgene ou Tempus)

La proposition de prix devra être forfaitaire par prestation (unité d'œuvre). Le Titulaire devra détailler les différentes offres de séquençage proposées par prestation.

Le CEA s'engage sur un montant de 150 k€ minimum sur la durée du marché.

DUREE DU MARCHÉ

La durée des prestations est prévue pour une période de 36 mois suivie d'une période optionnelle de 1 * 12 mois.

Début envisagé de la prestation : 2eme semestre 2025