### Plan de gestion des donnees de l’equipe sChneider-Maunoury

### 1. COLLECTE ET REUTILISATION DES DONNEES

### Informations sur la structure

**Nom de la structure**

Equipe “Développement du système nerveux central des Vertébrés”, unité Dev2A, IBPS

**Type de structure**

Equipe de recherche

**Identifiant de la structure**

Equipe SSM

**Responsabilités dans la structure**

Direction : Sylvie Schneider-Maunoury ([sylvie.schneider-maunoury@sorbonne-universite.fr](mailto:sylvie.schneider-maunoury@sorbonne-universite.fr) )

**Etablissement(s) tutelle(s)**

Sorbonne Université, CNRS (UMR 8263), Inserm (U1345).

**Projets** **et** **Financeur(s)** (*permettant l’acquisition des jeux de données – hors projet)*

1) Remodelage du lumen spinal (porteur: Pierre-Luc Bardet) – financements obtenus : Emergence SU 2021-2022 ; Gefluc 2025 ; Ligue Région IdF 2025-2026.

utilisatur.ice.s : Pierre-Luc Bardet, Amina Medyouf, Ana Maria Daza Zapata.

2) « NDCil » : Rôle des cils dans le développement du SNC et les ciliopathies (porteuse : Sylvie Schneider-Maunoury

Financement ERA-NET Neuron 2021 NDCil, ref. ANR-21-NEU2-0009-03.

3) « CiCerO » : rôle des cils primaires dans le cervelet (porteuse Sylvie Schneider-Maunoury)

Financement ANR PRC 2024

4) « Scoliose » : Rôle des cils primaires dans la scoliose chez le poisson zèbre (porteuse : Christine Vesque) – Financement de la Fondation Cotrel 2024

5) « BrainsToRM-112 » : Méthytransférases du complexe Trmt-112 dans les déficiences intellectuelles (porteuse : Sylvie Schneider-Maunoury) –

Financement ANR 2021 « BrainsToRM-112 »

Récurrents : Sorbonne U, CNRS, Inserm

### Informations sur le plan de gestion

**Historique des versions**

v0 – version martyr pour journée des plateformes – 4 mars 2025 (Pierre-Luc Bardet)

v1 – version initiale pour accès au serveur Psilo – 11 mars 2025 (Pierre-Luc Bardet)

v2 – version complétée le 11 mars 2025 (Sylvie Schneider-Maunoury)

### Informations sur lA COLLECTE DE DONNEES

**Mode d'obtention des données**

Données générées par la structure

**Type de données** *(entrée générique dans les templates de metadata : (INRAE, [DataCite](https://datacite-metadata-schema.readthedocs.io/en/4.6/properties/resourcetype/" \l "id1)...)*

* Images/Images de microscopie
* Jeu de données (Dataset)/Données de séquencage ARN

**Origine** *(Recommandations: L'origine des données correspond à la métadonnée Data Origin dans l'entrepôt* [*Data INRAE*](https://data.inrae.fr/)*)*

Données expérimentales

**Nature des données**

* Fichiers bruts d’acquisition sur microscopes de type confocal de la plateforme de l’IBPS.
* Fichiers bruts de RNAseq obtenus via la plateforme IGenSeq de l’ICM (Institute du cerveau).

**Format des données**

|  |  |  |  |
| --- | --- | --- | --- |
| ***Données d’imagerie*** | | | |
| **microscope** | **Logiciel d’acquisition** | **Format des fichiers bruts** | **Ouvrable avec** |
| Confocal Zeiss LSM 980 | ZEN microscopy software | .czi | ZEN, ImageJ, Fiji,… |
| Confocal Leica SP5 & SP8 | LAS software | .lif | LAS, ImageJ, Fiji,… |
| Yokogawa spinning disk | Metamorph | .nd, .tif | ImageJ, Fiji,… |
| ***Données de RNAseq*** | | | |
| **Séquenceur (Service)** | **Logiciel d’acquisition** | **Format des fichiers bruts** | **Ouvrable avec** |
| Novaseq6000 ILLUMINA, paired end. |  | fastqsanger.gz |  |
|  |  |  |  |

Tout effort pour convertir les données brutes d’imagerie en format ouvert (OME-TIFF, for more information : <https://docs.openmicroscopy.org/latest/ome-model/ome-tiff/>) SANS PERTE D’INFORMATION est la bienvenue.

**REUTILISATION DES DONNEES**

Les données seront accessibles en lecture par les autres membres de l’équipe, leur utilisation est possible en demandant à la personne qui a généré les données.

### 2. documentation ET Organisation des données

**Quels méthodes et outils sont utilisés pour acquérir et traiter les données, depuis leur acquisition jusqu'à leur mise à disposition, leur archivage ou leur destruction ?**

Données d’imagerie acquises sur un des microscopes de type confocal de la plateforme de l’IBPS. Au plus vite après leur acquisition, ces données brutes en sortie de microscope (ou converti sans perte en ome.tif), ou après traitement dans le cas de la technologie Airyscan, sont transférées sur le serveur de stockage. Ces données brutes sont conservées intactes, et TOUJOURS copiées sur un ordinateur de travail avant d’être manipulées.

Données de RNAseq acquises sur la plateforme IGenSeq de l’Institut du Cerveau (ICM, Paris, <https://institutducerveau.org/plateforme-sequencage-genomique-igenseq>). Les données brutes sont transférées via GLOBUS sur les serveurs GALAXY de la plateforme InforBio de l’IBPS ( ainsi que sur PSILO).

**Quelles métadonnées seront utilisées pour accompagner le jeu de données ?**

Pour les données d’imagerie, les métadonnées sont celles des différents logiciels propriétaires des microscopes utilisés. Dans le cas du format OME-TIF, il est possible d’AJOUTER des champs (utilisateur.ice, date, etc.) SANS supprimer les données fournies par le logiciel constructeur.

Pour les données de séquençage, les métadonnées sont celles fournies par la plateforme de séquençage, mises sous forme de tableau comme requis pour la soumission des données sur une plateforme de

**Comment les fichiers de données sont-ils gérés et organisés : contrôle des versions, conventions de nommage des fichiers, organisation des fichiers**

Les fichiers seront sauvegardés selon une architecture précise et à respecter SCRUPULEUSEMENT ! Il faudra les mettre dans un dossier par projet de l’équipe (voir la liste des projets dans la partie 1 ci-dessus). Puis un dossier au nom de la personne qui a produit ces données, qui renferment tous les dossiers de cette personne pour ce projet. Chaque expérience y est sauvegardée dans un dossier, dont le nom suit la **nomenclature** suivante :

Ex pour l’imagerie: 250301\_igu45hpf\_DAPI\_EdU\_HuC\_ZO-1

Ex pour le RNAseq : 250301\_Bulk\_WT\_

Soit, dans l’ordre pour l’imagerie:

- la **date**, en 6 chiffres, en commençant par l’année (25 pour 2025), puis le mois (03 pour mars) puis le jour (01 pour le 1er). Permet de classer par date en classant par ordre alphabétique.

- un **identifiant** qui donne une indication sur le contenu : mutant, traitement, stade de développement, etc. Laissé au choix de l’utilisateur-ice, les détails devant figurer dans le cahier de laboratoire à la date précisée plus haut.

- Le **type de marquage** réalisé, dans l’**ordre des longueurs d’onde** laser : DAPI en UV, marquage EdU en 488, révélation anticorps de 2 proteines en rouge (HuC) et rouge lointain (ZO-1).

Soit, dans l’ordre pour le RNAseq :

* La **date**, en 6 chiffres, en commençant par l’année (25 pour 2025), puis le mois (03 pour mars) puis le jour (01 pour le 1er). Permet de classer par date en classant par ordre alphabétique.
* Le **type d’expérience** : bulk, single cell
* Un **identifiant** qui donne une indication sur le contenu : type de matériel (org (organoïdes), m (mouse), zf (zebrafish)) et l’organe considéré : CB(cerebellum) , SC (spinal cord), CX (cortex)
* Les différents génotypes et stades, qui font partie d’une même expérience, seront indiqués dans le nom de chaque échantillon.

Les fichiers à l’intérieur du dossier sont organisés conformément à ce qui est noté dans le cahier de laboratoire.

**Quel est le processus de contrôle qualité des données ?**

La responsable de l’équipe se réserve le droit d’effacer toute donnée dont la sauvegarde n’aura pas respecté la nomenclature indiquée ci-dessus, et donc non-identifiable.

### 3. Stockage et sécurité des données

**Quels types de supports physiques sont utilisés pour stocker les données  ?**

Les données seront sauvegardées sur le serveur Psilo, conformément aux conditions décrites dans le PGD de la platerforme Inforbio à venir soon !

Une seconde copie sera conservée par les utilisateur-ices sur un disque-dur mis à disposition par l’équipe.

**Quelle est la volumétrie actuelle et prévisionnelle ?**

Imagerie : En se basant sur l’existant et les projets, de l’ordre de 1To par an minimum (à réviser).

RNAseq : En se basant sur l’existant et les projets, de l’ordre de 1To par an minimum (à réviser)

**4. ASPECTS ÉTHIQUES ET JURIDIQUES**

**Si des données à caractère personnel sont traitées, comment le respect des dispositions de la législation sur les données à caractère personnel et sur la sécurité des données sera-t-il assuré ?**

Aucune donnée personnelle n’est stockée sur Psilo.

### 5. PARTAGE ET Archivage des données

**Quelles sont les données à conserver sur le moyen ou le long terme et quelles sont les données à détruire ?**

Un point régulier des utilisateur-ices permettra d’effacer les expériences dont l’analyse aura révélé qu’elles n’ont pas contribué au projet. Le reste sera conservé.

Les données de RNAseq seront supprimées soit quand elles ont été publiées sur une plateforme d’archivage et rendues publiques (après publication), soit lorsqu’elles ne sont plus utilisées par l’équipe et donc conservées à long terme par les soins de l’équipe par un autre moyen.

**Sur quelle plateforme d'archivage pérenne seront archivées les données à conserver sur le long terme ? Sinon, quelles procédures seront mises en place pour la conservation à long terme ?**

A déterminer avec le service d’archivage de Sorbonne U.

**Quelle est la durée de conservation des données ?**

5 ans après la publication de l’article.

### 6. Responsabilités et ressources en matière de gestion des données

**Qui (par exemple rôle, position et institution de rattachement) sera responsable de la gestion des données (c'est-à-dire le gestionnaire des données) ?**

Chaque membre est responsable des données contenues dans le répertoire à son nom.

Lorsque la personne quitte l’’équipe, un autre membre de l’équipe reprend la responsabilité des données (à déterminera avec la responsable d’équipe : sylvie.schneider-maunoury@sorbonne-universite.fr

La responsable d’équipe contrôlera régulièrement la qualité et l’organisation de ces données.

Le PGD sera remis à jour tous les ans en cas de changement.

**Quelles garanties de financements couvriront les coûts associés à la conservation à long terme ?**

L’équipe a de nombreuses ressources propres. L’équipe dispose également de ressources de dotation des tutelles (CNRS, SU), qui permettent de garantir ces coûts même en cas (non envisageable dans les 4 prochaines années) d’absence de ressources propres.