

**Institute of Primate Research**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **SOP No.** | **Issue Number** | **Issue Date** | **Revision Status** | **Revision Date** |
| **SOP/KIPRE/RPD/DSAS/3.1.76** | **Version 01** | **October 2025** | **-** | **-** |

**STANDARD OPERATING PROCEDURE (SOP) DOCUMENT**

**Genome and Proteome Data Management**

|  |  |  |  |
| --- | --- | --- | --- |
| **Approvals** |  |  |  |
|  | **Name** | **Signature** | **Date** |
| **Developed by:** | \_Patrick Waweru Mwaura\_ | **\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_** | **\_6th October; 2025\_** |
|  | \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ | **\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_** | **\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_** |
|  | \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ | **\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_** | **\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_** |
| **Reviewed by:** | **\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_** | **\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_** | **\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_** |
| **Approved by:** | **\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_** | **\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_** | **\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_** |

|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  |  |

**Table of Contents**

[1. PURPOSE 4](#_Toc144316958)

[2. SCOPE 4](#_Toc144316959)

[3. PERSONS RESPONSIBLE: 4](#_Toc144316960)

[4. FREQUENCY 4](#_Toc144316961)

[5. MATERIALS 4](#_Toc144316962)

[6. PROCEDURE 4](#_Toc144316963)

[7. REFERENCES 7](#_Toc144316964)

# 

# PURPOSE

To provide standard procedures for secure, compliant, and reproducible management of genomic and proteomic datasets generated or curated by DS&AS.

# SCOPE

Applies to all genome and proteome data collected, processed, or stored under DS&AS, including raw sequence reads, processed alignments, annotations, and proteomic mass-spectrometry data.

# PERSONS RESPONSIBLE:

* **Bioinformatician:** Oversees data handling and quality control.
* **Data Engineer:** Maintains secure databases and repositories.
* **Head of DS&AS:** Ensures compliance with institutional and regulatory standards.

# FREQUENCY

* Continuous during project lifecycle.
* Annual data audits to ensure compliance.

# MATERIALS

* Secure servers/cloud storage with encryption.
* Reference databases (GenBank, UniProt, Ensembl).
* Metadata templates (MIAME/MIAPE standards).
* Institutional Data Protection and Sharing Policy.

# PROCEDURE

1. **Data Collection:** Store raw data in secure servers immediately after sequencing/analysis.
2. **Metadata Capture:** Document experimental details using MIAME/MIAPE standards.
3. **Quality Control:** Run sequence/proteome QC (FastQC, ProteoQC).
4. **Access Control:** Restrict access to sensitive datasets via role-based permissions.
5. **Archiving:** Maintain backups and mirror copies in long-term storage.
6. **Compliance:** Ensure anonymisation of sensitive human-derived data before sharing.

# REFERENCES