

# **Institute of Primate Research**

# STANDARD OPERATING PROCEDURE (SOP) DOCUMENT

## **Genome and Proteome Data Management**

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

Approvals			
	Name	Signature	Date
Developed by:	_Patrick Waweru Mwaura_		
Reviewed by: Approved by:			

# Table of Contents 4 1. PURPOSE 4 2. SCOPE 4 3. PERSONS RESPONSIBLE: 4 4. FREQUENCY 4 5. MATERIALS 4 6. PROCEDURE 4 7. REFERENCES 4

### 1. PURPOSE

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

### 2. 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.

### 3. 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.

### 4. FREQUENCY

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

### 5. MATERIALS

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

### 6. 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.

### 7. REFERENCES