MCP Servers Reference Guide

Comprehensive Bioinformatics Server Capabilities

Version 1.0 | Generated: December 10, 2025

This reference guide documents all Model Context Protocol (MCP) servers available for bioinformatics analysis. Each server provides specialized functionality for different aspects of genomic, transcriptomic, proteomic, and clinical data analysis. The servers work together to enable comprehensive multi-modal analysis workflows through standardized interfaces.

Table of Contents

1. mcp-pubmed

■ Biomedical Literature Search and Retrieval

Overview

The PubMed MCP server provides comprehensive access to biomedical and life sciences literature through the PubMed database. It enables searching, retrieving, and analyzing scientific publications with support for full-text access through PubMed Central.

Key Functions

Function	Description	Use Case
search_articles	Search PubMed with keywords or queries	Find relevant research papers
get_article_metadata	Retrieve detailed article information	Get authors, abstract, citations
convert_article_ids	Convert between PMID, PMCID, DOI	Cross-reference articles
find_related_articles	Find similar or related papers	Expand literature review
get_full_text_article	Retrieve full text from PMC	Access complete paper content
get_copyright_status	Check article licensing	Determine reuse permissions

Example Usage

- Search for BRCA1 mutations in ovarian cancer papers published in last 2 years
- Retrieve full text of high-impact papers for systematic review
- Find all papers by a specific author with their collaborators
- Check copyright status before reproducing figures

Limitations

Only indexes biomedical literature. Does not include physics, mathematics, pure chemistry, or engineering papers unless they have biomedical relevance.

2. mcp-fgbio

■ Genomic Data Processing and Quality Control

Overview

The FGBio MCP server provides tools for processing genomic sequencing data, including FASTQ quality control, UMI extraction, and gene annotation queries. It supports various reference genomes and annotation databases.

Key Functions

Function	Description	Use Case
fetch_reference_genome	Download reference genomes (hg38, mm10)	Alignment preparation
validate_fastq	QC validation of FASTQ files	Check sequencing quality
extract_umis	Extract Unique Molecular Identifiers	Prepare for deduplication
query_gene_annotations	Get gene annotations from GENCODE/Ensembl	Gene information lookup

Supported Genomes

- Human: hg38, hg19
- Mouse: mm10, mm39
- Rat: rn6
- Zebrafish: danRer11

3. mcp-huggingface

■ Genomic Foundation Models and AI

Overview

The HuggingFace MCP server provides access to state-of-the-art genomic language models for DNA, RNA, and protein sequence analysis. It enables cell type prediction, sequence embedding generation, and variant effect prediction using transformer-based models.

Key Functions

Function	Description	Use Case
load_genomic_model	Load pre-trained models (DNABERT, Geneformer)	Model initialization
predict_cell_type	Classify cell types from expression	Single-cell analysis
embed_sequences	Generate embeddings for sequences	Similarity analysis

Available Models

- DNABERT-2: DNA sequence analysis
- Geneformer: Single-cell transcriptomics
- Nucleotide-Transformer: Multi-species genomics
- scGPT/scBERT: Single-cell foundation models

4. mcp-seqera

■ Nextflow Pipeline Execution Platform

Overview

The Seqera MCP server enables execution of Nextflow bioinformatics pipelines through the Seqera Platform. It provides access to nf-core pipelines and custom workflows with cloud compute integration.

Key Functions

Function	Description	Use Case
launch_nextflow_pipeline	Execute nf-core or custom pipelines	Run analysis workflows
monitor_workflow_status	Track pipeline execution progress	Monitor job status
list_available_pipelines	Query available pipeline catalog	Discover workflows

Supported Pipelines

- nf-core/rnaseq: RNA sequencing analysis
- nf-core/sarek: Variant calling
- nf-core/spatial: Spatial transcriptomics
- nf-core/chipseq: ChIP-seq analysis
- Custom user pipelines

5. mcp-mockepic

■ Electronic Health Records (Mock/Demo)

Overview

The MockEpic MCP server simulates an Epic-like electronic health record system for demonstration purposes. It provides access to patient demographics, clinical data, laboratory results, and medication histories.

Key Functions

Function	Description	Use Case
query_patient_records	Retrieve patient demographics and history	Get clinical context
link_spatial_to_clinical	Connect omics to clinical data	Integrate multi-modal data
search_diagnoses	Query ICD-10 diagnosis codes	Find disease codes

Data Types

- Patient demographics (age, sex, ethnicity)
- Laboratory results (CBC, chemistry panels, tumor markers)
- Medication history
- Diagnosis codes (ICD-10)
- Clinical notes (simulated)

6. mcp-deepcell

■ Deep Learning Cell Segmentation

Overview

The DeepCell MCP server provides deep learning-based cell segmentation for microscopy images. It can identify and segment individual cells, classify cell states, and analyze cellular phenotypes from various imaging modalities.

Key Functions

Function	Description	Use Case
segment_cells	AI-based cell segmentation	Identify individual cells
classify_cell_states	Determine cell states/phenotypes	Cell type classification

Segmentation Models

- Nuclear segmentation: For DAPI/Hoechst stained nuclei
- Membrane segmentation: For cell membrane markers
- Cytoplasm segmentation: For cytoplasmic stains
- Multi-channel segmentation: Combined nuclear/membrane

7. mcp-openimagedata

■■ Histology and Imaging Data Management

Overview

The OpenImageData MCP server manages histology and microscopy images, including H&E; stains and immunofluorescence. It provides image registration, feature extraction, and integration with spatial transcriptomics data.

Key Functions

Function	Description	Use Case
fetch_histology_image	Retrieve H&E or IF images	Access tissue images
register_image_to_spatial	Align images with spatial data	Spatial integration
extract_image_features	Extract morphological features	Quantitative analysis

Supported Image Types

- H&E; (Hematoxylin & Eosin) stained slides
- Immunofluorescence (IF) images
- Immunohistochemistry (IHC) images
- Multiplex immunofluorescence
- Whole slide images (WSI)

8. mcp-tcga

■■ The Cancer Genome Atlas Database

Overview

The TCGA MCP server provides access to The Cancer Genome Atlas, containing genomic, transcriptomic, and clinical data for over 30 cancer types. It enables comparison of patient samples with large cohorts for molecular subtyping and prognosis.

Key Functions

Function	Description	Use Case
query_tcga_cohorts	Search cancer cohorts by type	Find relevant datasets
fetch_expression_data	Get gene expression profiles	Expression analysis
compare_to_cohort	Statistical comparison with TCGA	Patient stratification
get_survival_data	Retrieve survival correlations	Prognostic analysis
get_mutation_data	Query mutation frequencies	Mutation landscape

Cancer Types (Examples)

- BRCA: Breast invasive carcinoma
- OV: Ovarian serous cystadenocarcinoma
- LUAD: Lung adenocarcinoma
- COAD: Colon adenocarcinoma
- GBM: Glioblastoma multiforme

9. mcp-spatialtools

■ Spatial Transcriptomics Analysis

Overview

The SpatialTools MCP server provides comprehensive analysis tools for spatial transcriptomics data from platforms like 10X Visium, Xenium, and MERFISH. It handles QC, spatial statistics, differential expression, and integration with histology images.

Key Functions

Function	Description	Use Case
filter_quality	QC filtering of spatial spots	Data preprocessing
split_by_region	Segment by spatial regions	Region-specific analysis
align_spatial_data	Align reads with STAR	Read mapping
merge_tiles	Combine multiple tissue sections	Multi-section integration
calculate_spatial_autocorrelation	Moran's I statistics	Spatial patterns
perform_differential_expression	Compare regions/conditions	Find markers
perform_batch_correction	Correct batch effects	Data harmonization
perform_pathway_enrichment	Pathway analysis	Biological interpretation

Supported Platforms

- 10X Genomics Visium
- 10X Genomics Xenium
- MERFISH
- Slide-seq
- Generic spatial formats

10. mcp-multiomics

■ Multi-Omics Data Integration

Overview

The MultiOmics MCP server integrates data across multiple molecular modalities including RNA-seq, proteomics, phosphoproteomics, and metabolomics. It provides statistical methods for cross-modality analysis and meta-analysis.

Key Functions

Function	Description	Use Case
integrate_omics_data	Align and normalize multi-omics	Data integration
run_halla_analysis	All-against-all associations	Find correlations
calculate_stouffer_meta	Combine p-values across modalities	Meta-analysis
create_multiomics_heatmap	Integrated visualization	Data visualization
run_multiomics_pca	Dimensionality reduction	Sample clustering

Supported Data Types

- Transcriptomics: RNA-seq, microarray
- Proteomics: Mass spectrometry, TMT/iTRAQ
- Phosphoproteomics: Phospho-enriched MS
- Metabolomics: LC-MS, NMR
- Epigenomics: DNA methylation, ChIP-seq

Server Integration and Workflows

Common Workflow Patterns

The MCP servers are designed to work together in integrated workflows. Here are common analysis patterns that combine multiple servers:

Workflow	Servers Used	Description
Literature-guided analysis	pubmed → tcga → multiomics	Find relevant papers, compare with TCGA, integrate data
Spatial transcriptomics pipeline	fgbio → spatialtools → openimagedata	Process reads, analyze spatial data, integrate histology
Clinical genomics	mockepic → tcga → multiomics	Get clinical data, compare with TCGA, integrate omics
Image-guided analysis	openimagedata → deepcell → spatialtools	Process images, segment cells, map to spatial data
Pipeline automation	seqera → fgbio → spatialtools	Run Nextflow pipelines for complete analysis
AI-enhanced analysis	huggingface → multiomics → tcga	Use AI models, integrate results, compare cohorts

Server Categories

MCP servers can be grouped into functional categories based on their primary role in the analysis workflow:

Category	Servers	Primary Role
Data Processing	fgbio, seqera	Raw data processing, QC, pipelines
Clinical Integration	mockepic, tcga	Clinical context, cohort comparison
Imaging & Spatial	deepcell, openimagedata, spatialtools	Image analysis, spatial data
Analysis & Integration	multiomics, huggingface	Data integration, AI analysis
Knowledge & Literature	pubmed	Scientific literature access

Best Practices and Guidelines

Server Selection Guidelines

- Choose servers based on data type: Use spatialtools for Visium, deepcell for microscopy
- Start with quality control: Always run QC (fgbio, spatialtools) before analysis
- Validate with external data: Compare results with TCGA or literature (pubmed)
- Integrate incrementally: Start with single modalities before multi-omics integration
- Use appropriate models: Select HuggingFace models matched to your data type
- Document workflows: Track which servers and functions were used for reproducibility

Performance Considerations

- Batch operations when possible to reduce API calls
- Cache results from expensive operations (TCGA queries, model predictions)
- Use appropriate data formats (compressed FASTQ, indexed BAM)
- Consider compute requirements for pipeline execution (seqera)
- Optimize image resolution for analysis needs

Data Privacy and Security

- MockEpic is for demonstration only - use secure EHR integration in production
- Ensure HIPAA compliance when processing patient data
- Use appropriate consent for genomic data analysis
- Implement access controls for sensitive data
- Maintain audit logs for clinical data access

Quick Reference Guide

Task	Primary Server	Alternative/Supporting
Search literature	mcp-pubmed	-
Process FASTQ files	mcp-fgbio	mcp-seqera
Run RNA-seq pipeline	mcp-seqera	mcp-fgbio
Analyze spatial data	mcp-spatialtools	mcp-openimagedata
Segment cells in images	mcp-deepcell	-
Compare with TCGA	mcp-tcga	-
Integrate multi-omics	mcp-multiomics	-
Get clinical data	mcp-mockepic	-
Use AI models	mcp-huggingface	-
Align histology images	mcp-openimagedata	mcp-spatialtools

Notes and Version Information

Document Version: 1.0
Last Updated: December 10, 2025
Total Servers Documented: 10

Additional Resources:
• MCP Protocol Documentation: https://github.com/modelcontextprotocol
• Server-specific documentation available in each server repository
• Example workflows and tutorials in the MCP examples repository

Future Additions:
• Additional genomic databases (GEO, SRA, dbGaP)
• Proteomics-specific servers (PRIDE, MassIVE)
• Clinical trial matching servers
• Drug-target interaction databases
• Pathway and network analysis servers

Contact:
For questions or suggestions about MCP servers, consult the official documentation or community forums.