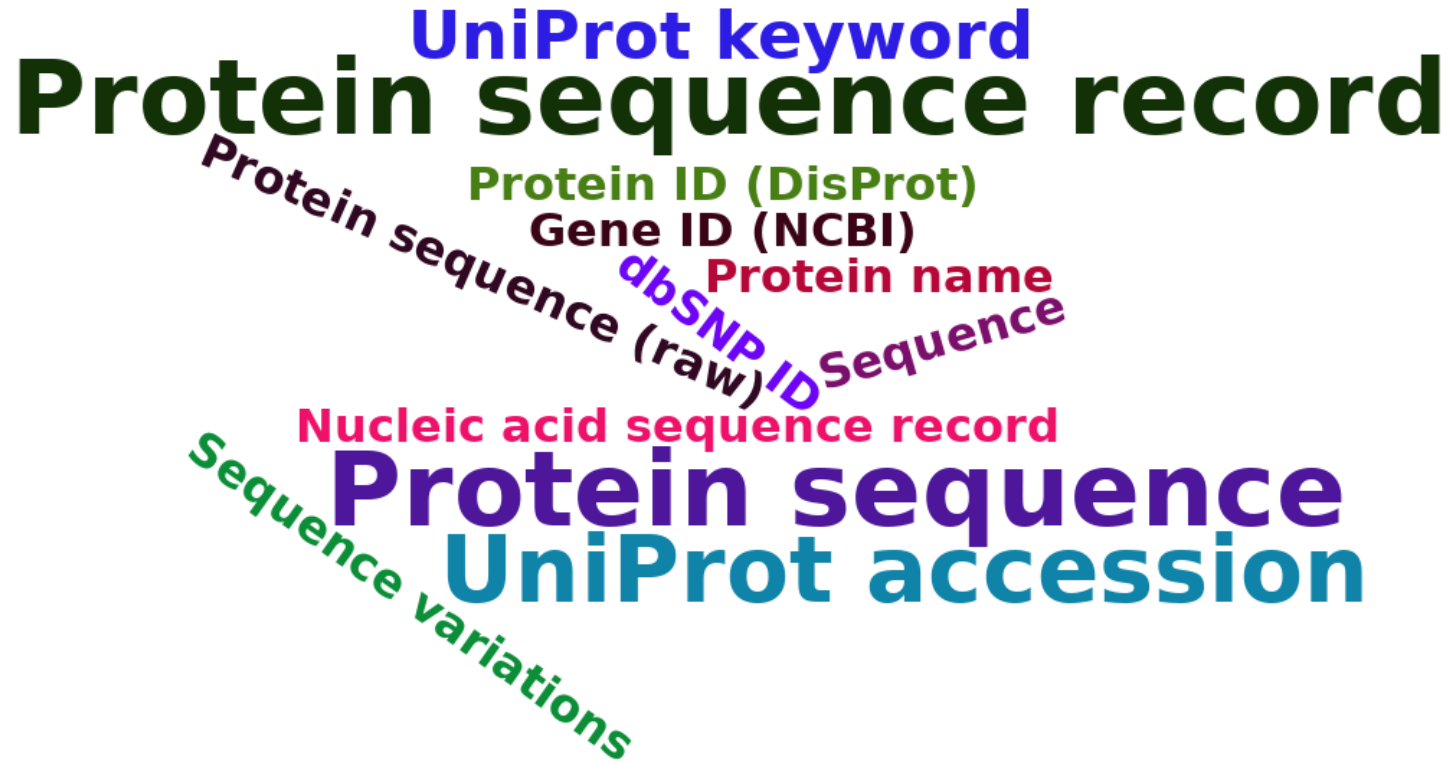
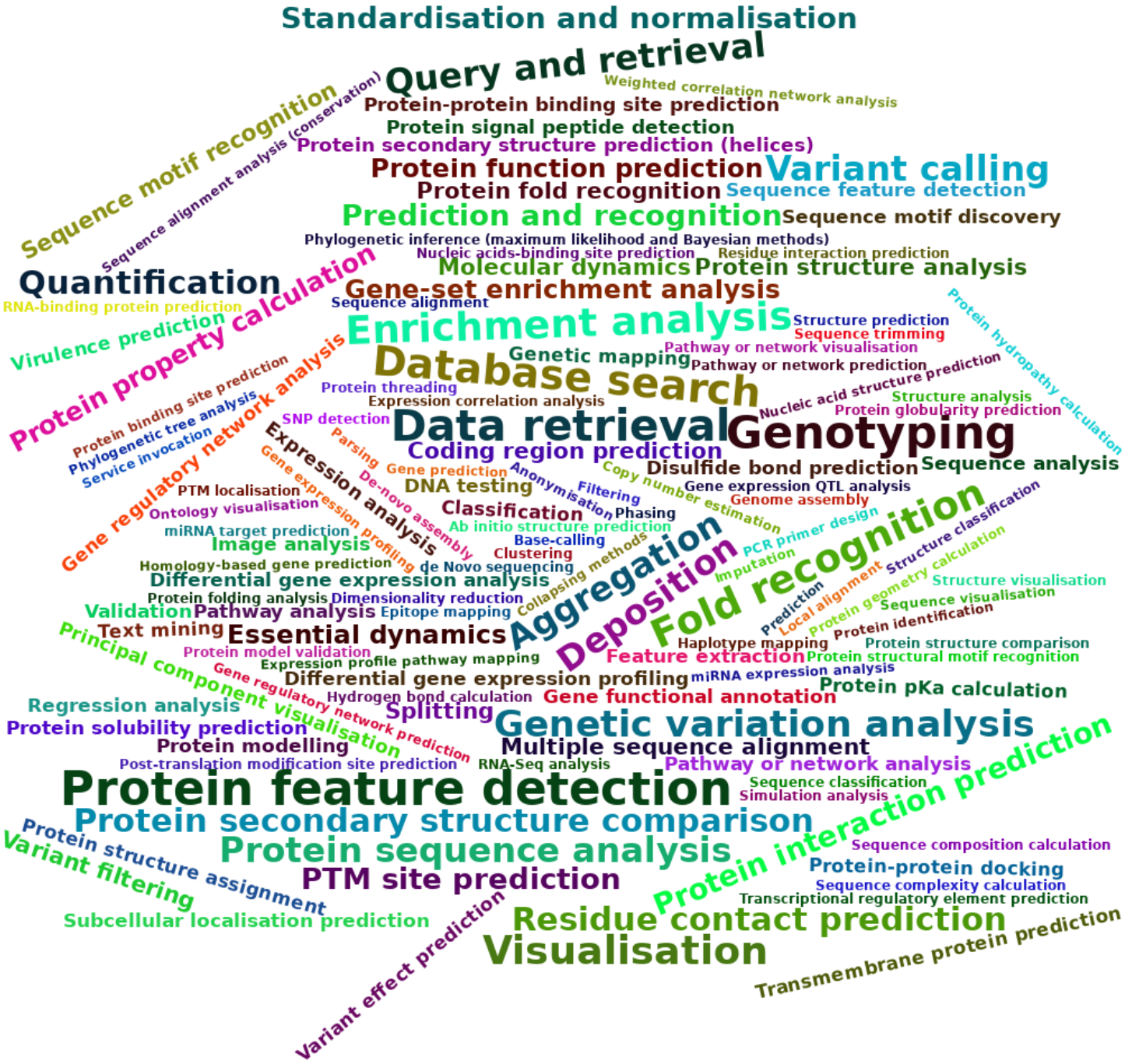

UniProt keyword
Protein sequence record
Protein ID (DisProt)
Protein sequence (raw)
Gene ID (NCBI)
dbSNP ID
Protein name
Sequence
Nucleic acid sequence record
Protein sequence
UniProt accession
Sequence variations

Textual format
FASTA
VCF

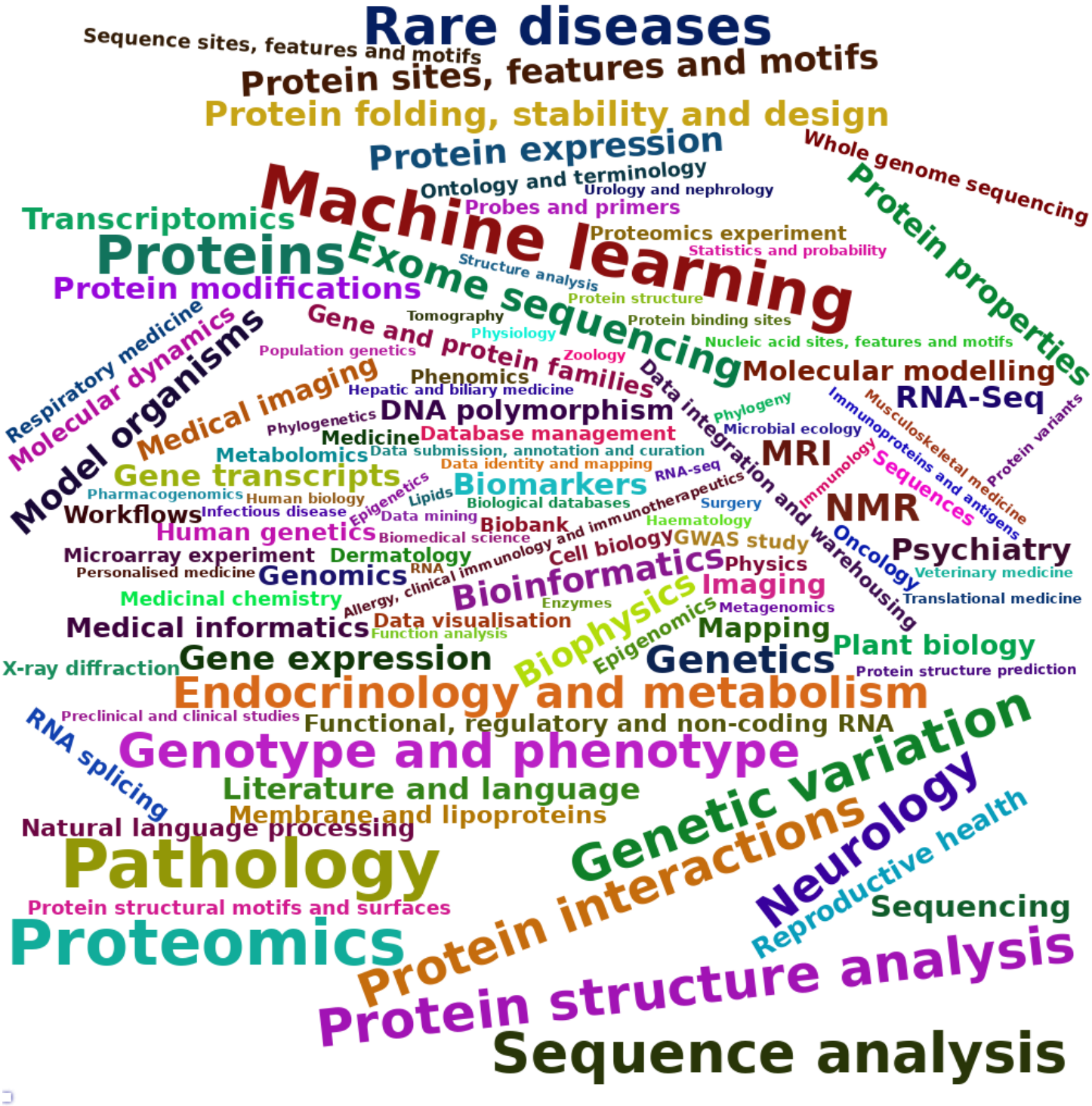
Unlicense
Other
CC-BY-NC-ND-4.0
ODbL-1.0
CC-BY-4.0
MIT
Not licensed
BSD-3-Clause
CC-BY-NC-1.0
CC-BY-NC-SA-4.0
GPL-2.0
GPL-3.0
Proprietary
CC-BY-NC-4.0

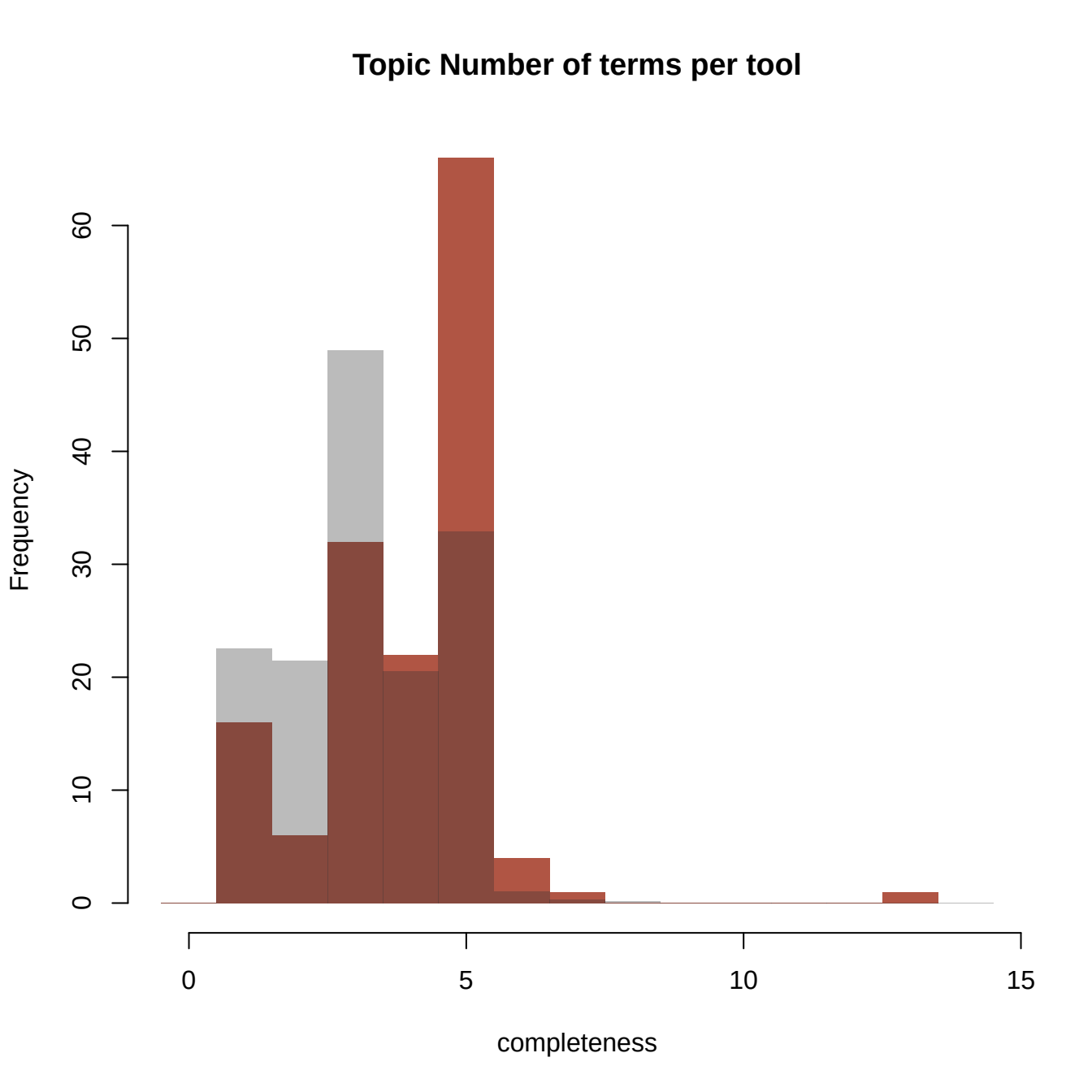

Standardisation and normalisation
Query and retrieval
Weighted correlation network analysis
Sequence motif recognition
Sequence alignment analysis (conservation)
Protein-protein binding site prediction
Protein signal peptide detection
Protein secondary structure prediction (helices)
Protein function prediction
Variant calling
Protein fold recognition
Sequence feature detection
Prediction and recognition
Sequence motif discovery
Phylogenetic inference (maximum likelihood and Bayesian methods)
Nucleic acids-binding site prediction
Residue interaction prediction
Molecular dynamics
Protein structure analysis
Quantification
Gene-set enrichment analysis
RNA-binding protein prediction
Sequence alignment
Protein property calculation
Virulence prediction
Enrichment analysis
Structure prediction
Sequence trimming
Protein hydropathy calculation
Pathway or network visualisation
Genetic mapping
Pathway or network prediction
Database search
Protein binding site prediction
Phylogenetic tree analysis
Service invocation
Protein threading
Expression correlation analysis
Nucleic acid structure prediction
Structure analysis
Protein globularity prediction
Gene regulatory network analysis
SNP detection
Data retrieval
Genotyping
Coding region prediction
Expression analysis
Parsing
Gene prediction
Anonymisation
Filtering
Copy number estimation
Disulfide bond prediction
Sequence analysis
Gene expression QTL analysis
Genome assembly
PTM localisation
Ontology visualisation
De-novo assembly
DNA testing
Classification
Phasing
Ab initio structure prediction
Gene expression profiling
miRNA target prediction
Image analysis
Base-calling
Clustering
Collapsing methods
PCR primer design
Imputation
Fold recognition
Structure classification
Homology-based gene prediction
de Novo sequencing
Differential gene expression analysis
Aggregation
Deposition
Structure visualisation
Protein geometry calculation
Local alignment
Prediction
Sequence visualisation
Protein identification
Protein folding analysis
Dimensionality reduction
Validation
Pathway analysis
Epitope mapping
Text mining
Essential dynamics
Haplotype mapping
Protein structure comparison
Principal component visualisation
Protein model validation
Expression profile pathway mapping
Feature extraction
Protein structural motif recognition
Gene regulatory network prediction
Differential gene expression profiling
miRNA expression analysis
Protein pKa calculation
Hydrogen bond calculation
Gene functional annotation
Regression analysis
Splitting
Genetic variation analysis
Protein solubility prediction
Protein modelling
Multiple sequence alignment
Post-translation modification site prediction
RNA-Seq analysis
Pathway or network analysis
Protein interaction prediction
Protein feature detection
Sequence classification
Simulation analysis
Protein secondary structure comparison
Protein structure assignment
Variant filtering
Protein sequence analysis
Sequence composition calculation
Protein-protein docking
PTM site prediction
Sequence complexity calculation
Transcriptional regulatory element prediction
Subcellular localisation prediction
Variant effect prediction
Residue contact prediction
Visualisation
Transmembrane protein prediction

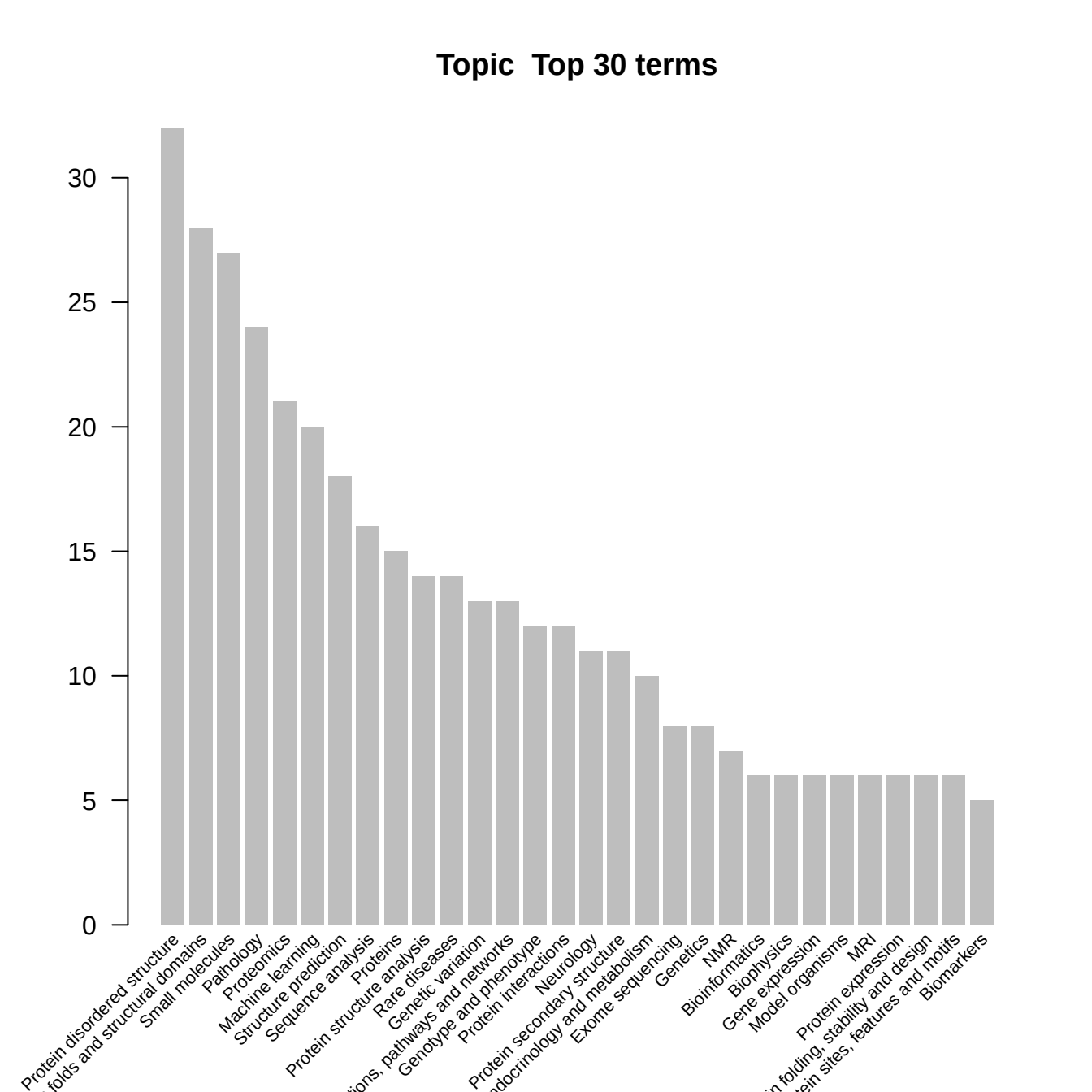
Sequence features
Protein structure report
Protein features
Feature table
Database search results
Profile-profile alignment
Protein structure
Protein report
Protein solubility
Sequence alignment

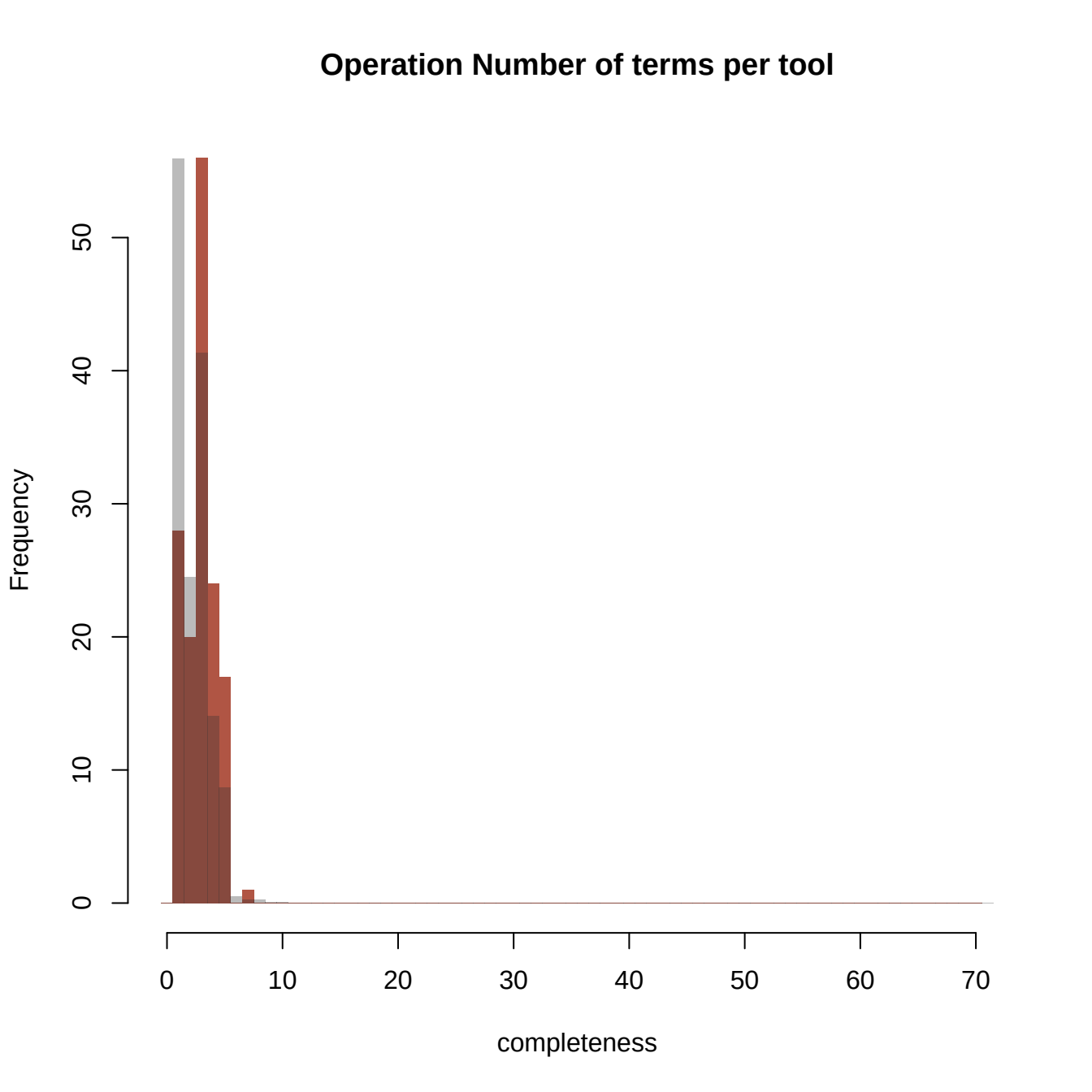
Sequence feature table format
JSON
TSV
FASTA
png
CSV
PDF
HTML
RDF format
plain text format (unformatted)

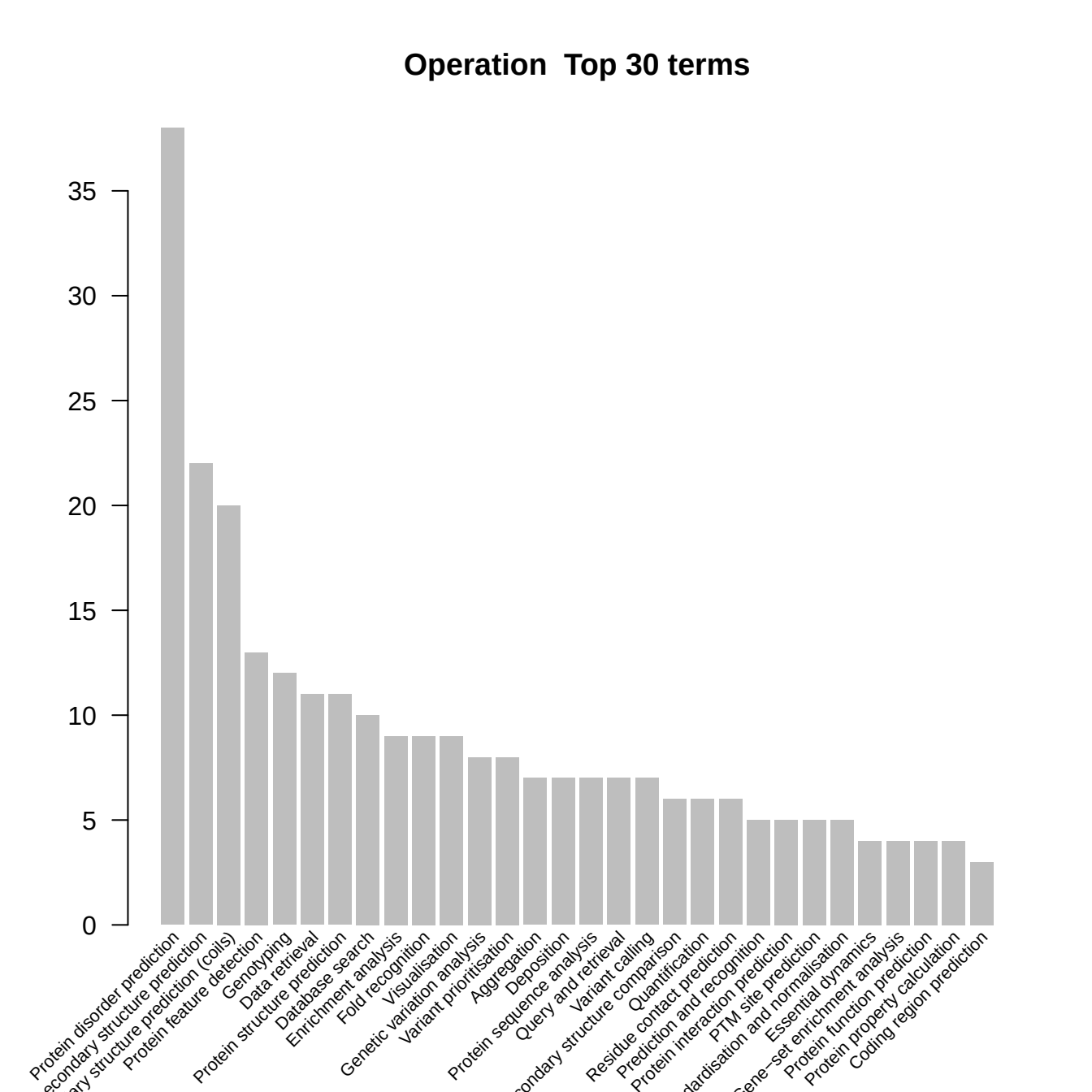
Web application
Web API
Desktop application
Workbench
SPARQL endpoint
Plug-in
Script
Command-line tool
Library
Web service
Suite
Database portal

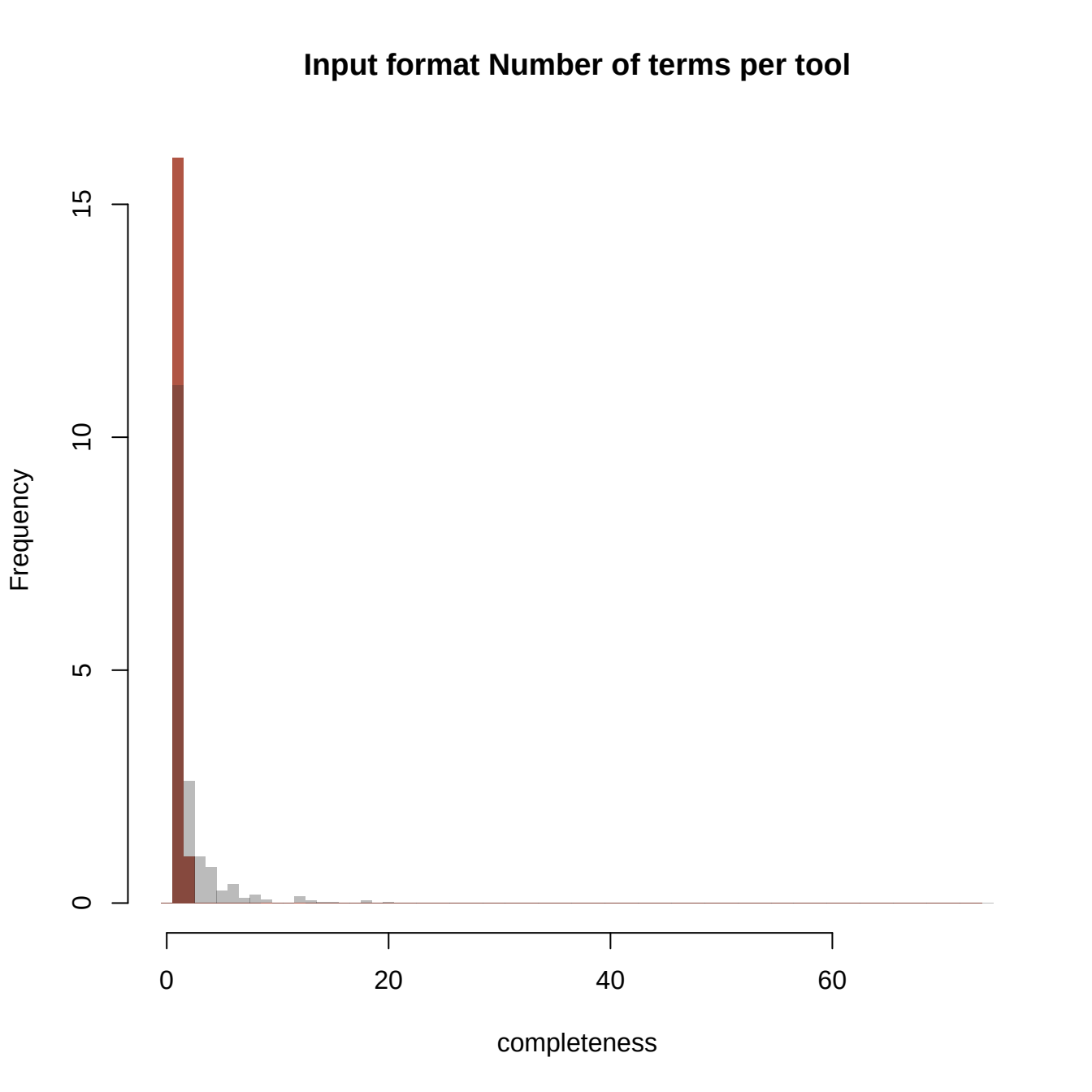

Rare diseases
Sequence sites, features and motifs
Protein sites, features and motifs
Protein folding, stability and design
Protein expression
Whole genome sequencing
Ontology and terminology
Urology and nephrology
Machine learning
Protein properties
Probes and primers
Transcriptomics
Proteomics experiment
Proteins
Statistics and probability
Exome sequencing
Structure analysis
Protein modifications
Protein structure
Tomography
Protein binding sites
Respiratory medicine
Molecular dynamics
Physiology
Nucleic acid sites, features and motifs
Gene and protein families
Population genetics
Zoology
Model organisms
Molecular modelling
Medical imaging
Phenomics
Data integration and warehousing
Hepatic and biliary medicine
RNA-Seq
Phylogeny
Immunoproteins and antigens
Musculoskeletal medicine
Protein variants
Phylogenetics
DNA polymorphism
Microbial ecology
Medicine
Database management
MRI
Metabolomics
Data submission, annotation and curation
Immunology
Sequences
Gene transcripts
Data identity and mapping
RNA-seq
Pharmacogenomics
Human biology
Biomarkers
Epigenetics
Lipids
Biological databases
Surgery
NMR
Workflows
Infectious disease
Data mining
Haematology
Biobank
Human genetics
Biomedical science
Oncology
Cell biology
GWAS study
Psychiatry
Microarray experiment
Dermatology
Bioinformatics
Physics
Personalised medicine
Genomics
RNA
Veterinary medicine
Allergy, clinical immunology and immunotherapeutics
Imaging
Medicinal chemistry
Translational medicine
Enzymes
Biophysics
Metagenomics
Medical informatics
Data visualisation
Mapping
Function analysis
Epigenomics
Plant biology
Gene expression
Genetics
X-ray diffraction
Protein structure prediction
Endocrinology and metabolism
Preclinical and clinical studies
Functional, regulatory and non-coding RNA
Genetic variation
RNA splicing
Genotype and phenotype
Literature and language
Neurology
Membrane and lipoproteins
Natural language processing
Reproductive health
Pathology
Protein interactions
Protein structural motifs and surfaces
Sequencing
Proteomics
Protein structure analysis
Sequence analysis

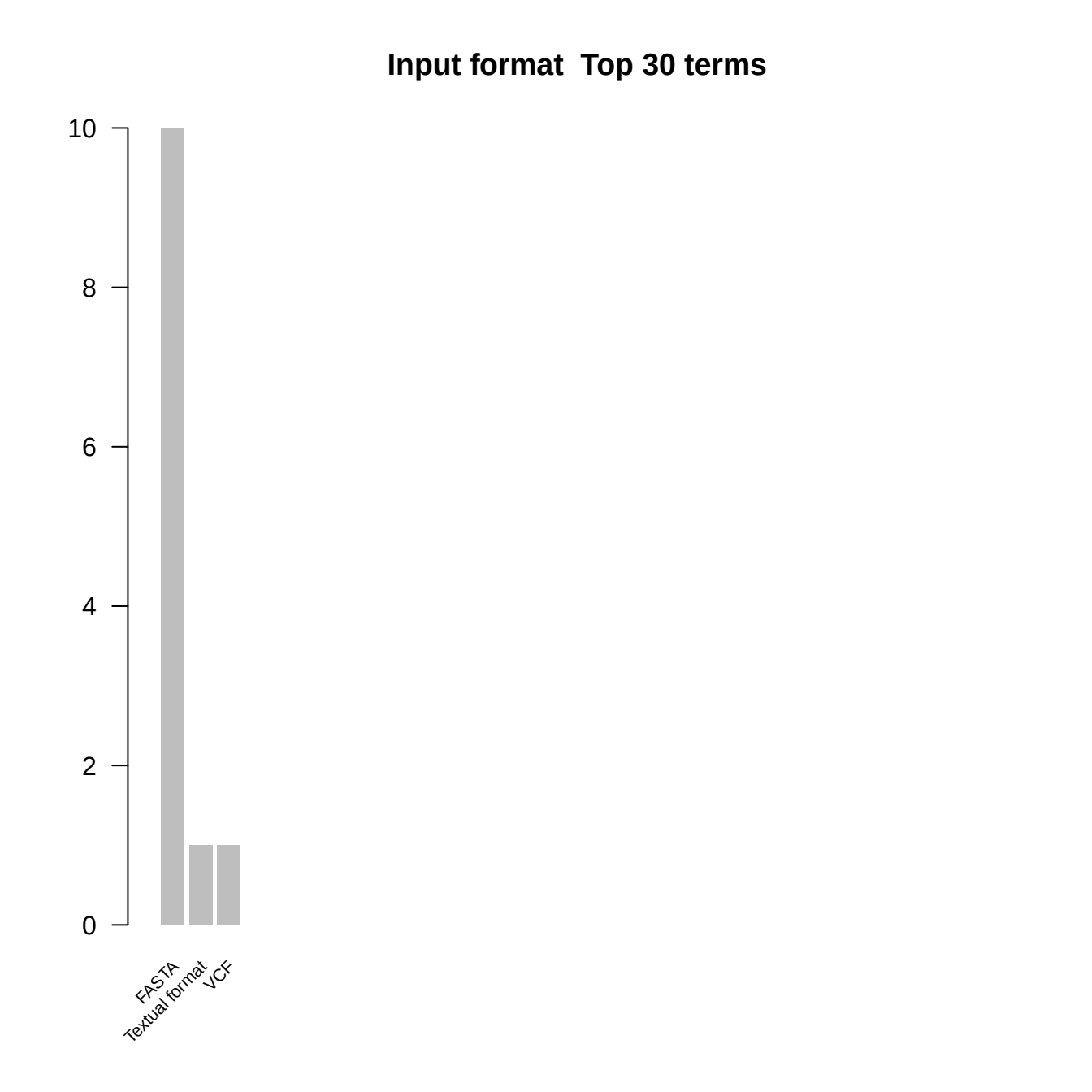

Topic Number of terms per tool
Frequency
0
10
20
30
40
50
60
0
5
10
15
completeness

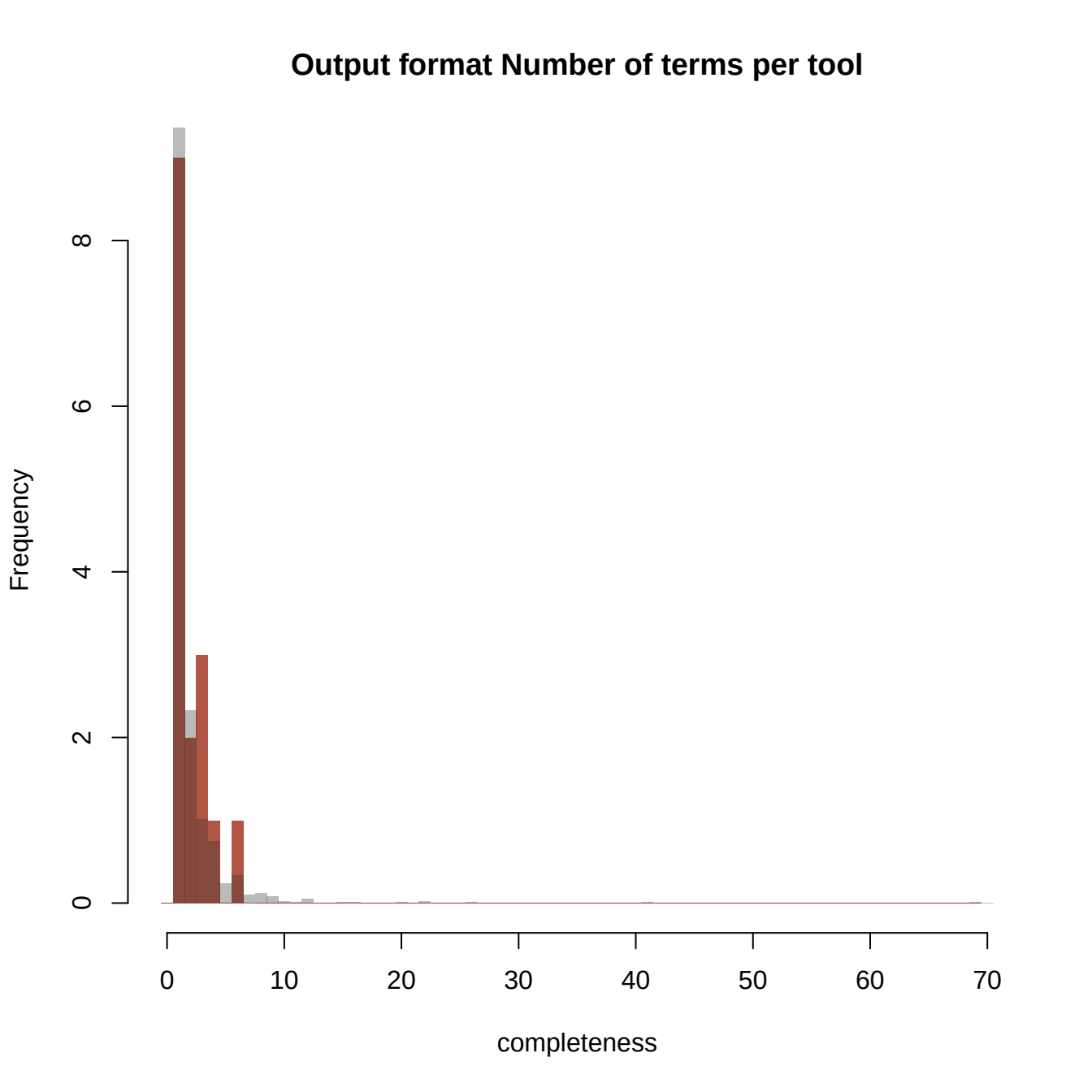

Topic Top 30 terms
30
25
20
15
10
5
0
Protein disordered structure
folds and structural domains
Small molecules
Pathology
Proteomics
Machine learning
Structure prediction
Sequence analysis
Proteins
Protein structure analysis
Rare diseases
Genetic variation
ions, pathways and networks
Genotype and phenotype
Protein interactions
Neurology
Protein secondary structure
ndocrinology and metabolism
Exome sequencing
Genetics
NMR
Bioinformatics
Biophysics
Gene expression
Model organisms
MRI
Protein expression
n folding, stability and design
tein sites, features and motifs
Biomarkers

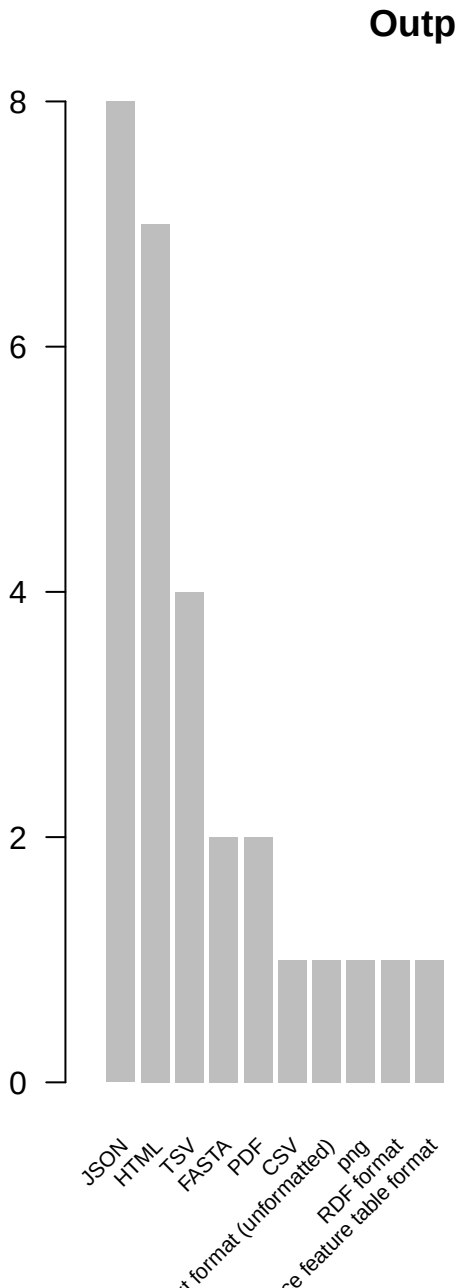

Operation Number of terms per tool
Frequency
0
10
20
30
40
50
0
10
20
30
40
50
60
70
completeness

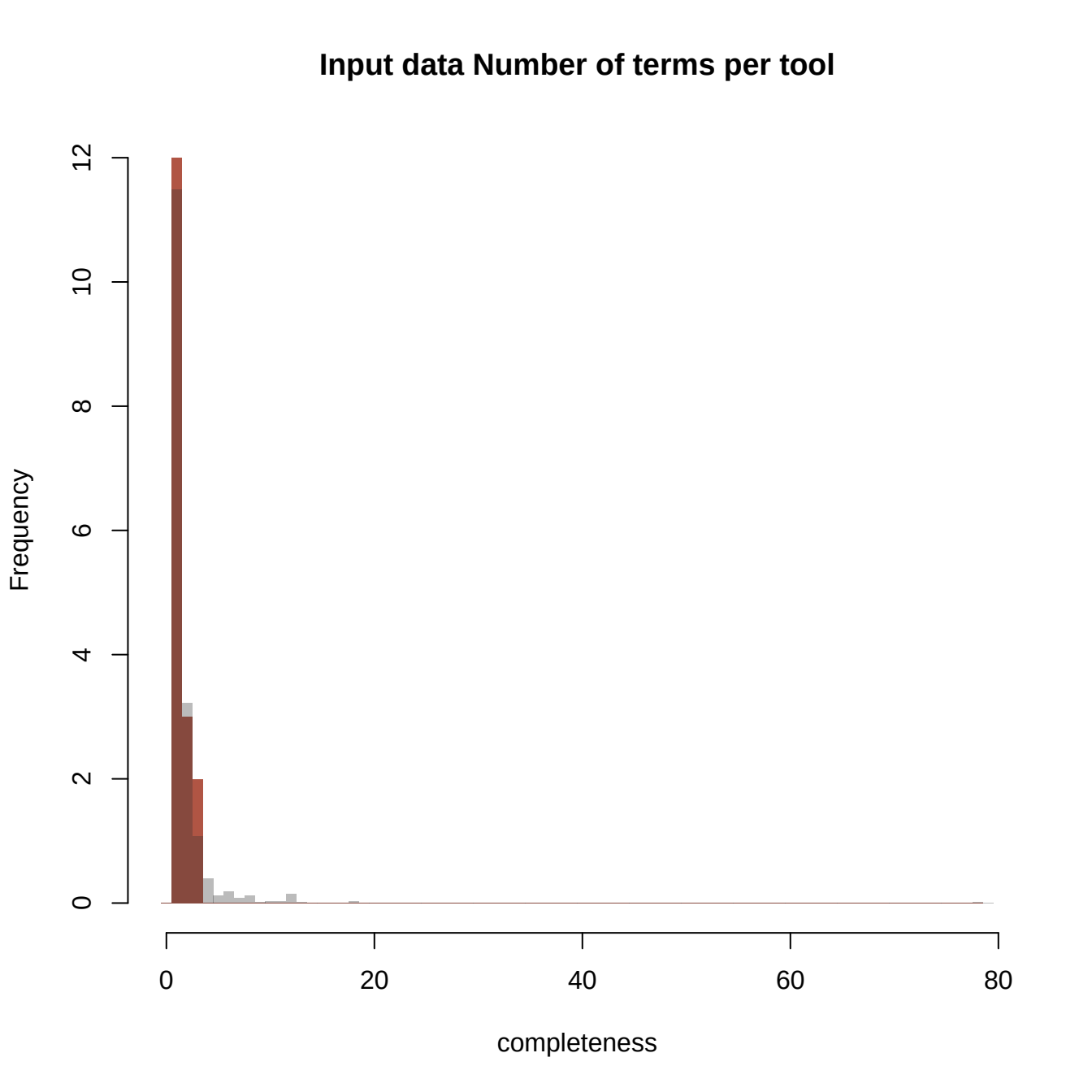

Operation Top 30 terms
0
5
10
15
20
25
30
35
Protein disorder prediction
Secondary structure prediction
Secondary structure prediction (coils)
Protein feature detection
Genotyping
Data retrieval
Protein structure prediction
Database search
Enrichment analysis
Fold recognition
Visualisation
Genetic variation analysis
Variant prioritisation
Aggregation
Deposition
Protein sequence analysis
Query and retrieval
Variant calling
Structure comparison
Quantification
Residue contact prediction
Prediction and recognition
Protein interaction prediction
PTM site prediction
Standardisation and normalisation
Essential dynamics
Gene-set enrichment analysis
Protein function prediction
Protein property calculation
Coding region prediction

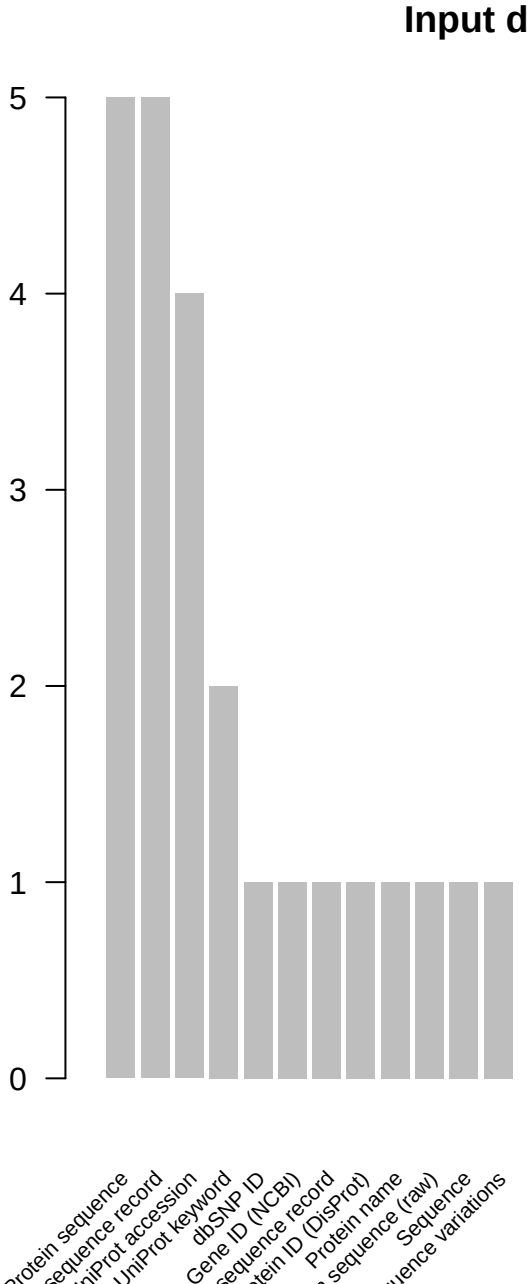

Input format Number of terms per tool
Frequency
0
5
10
15
0
20
40
60
completeness

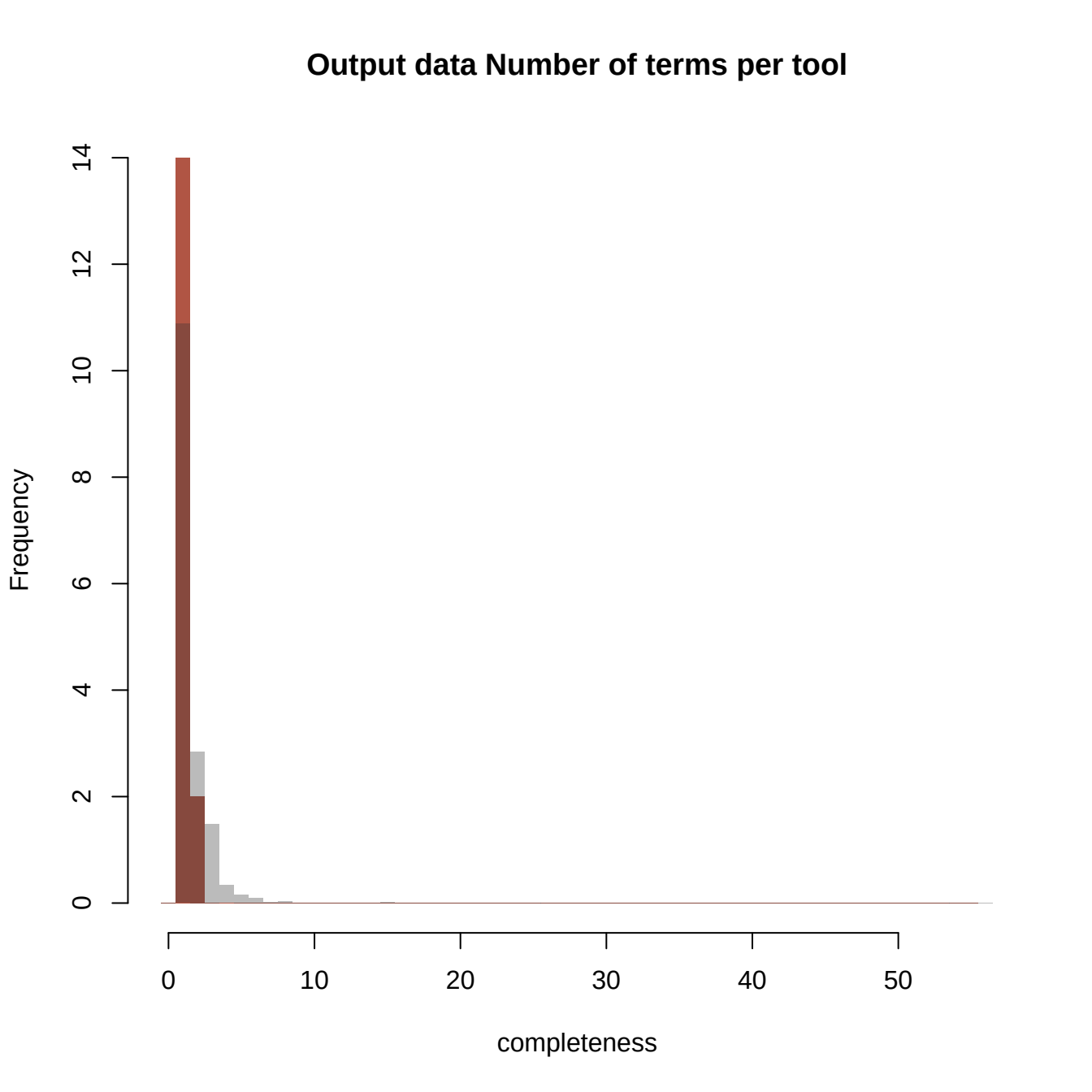

Input format Top 30 terms
10
8
6
4
2
0
FASTA
Textual format
VCF

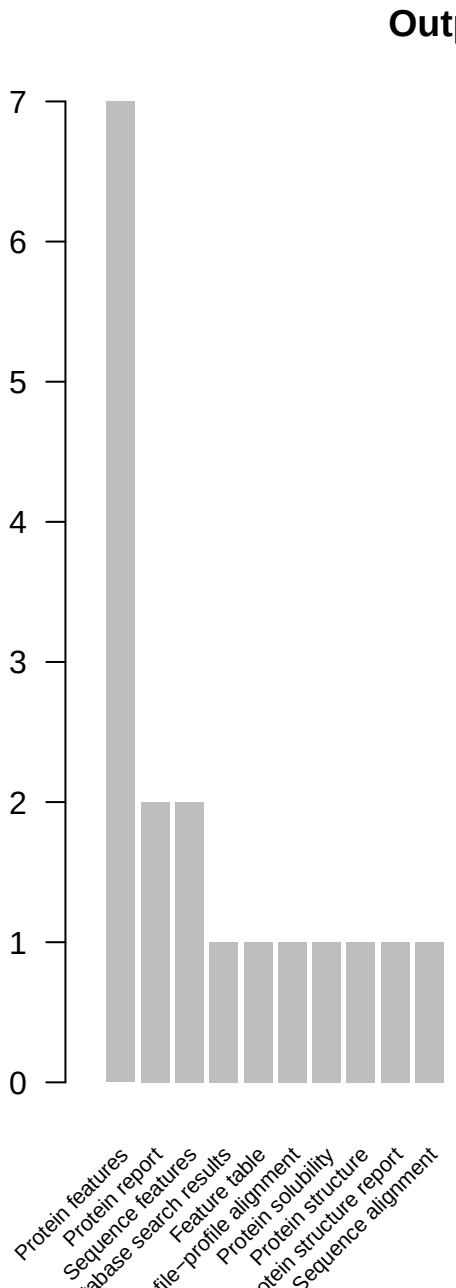

Output format Number of terms per tool
Frequency
0
2
4
6
8
0
10
20
30
40
50
60
70
completeness

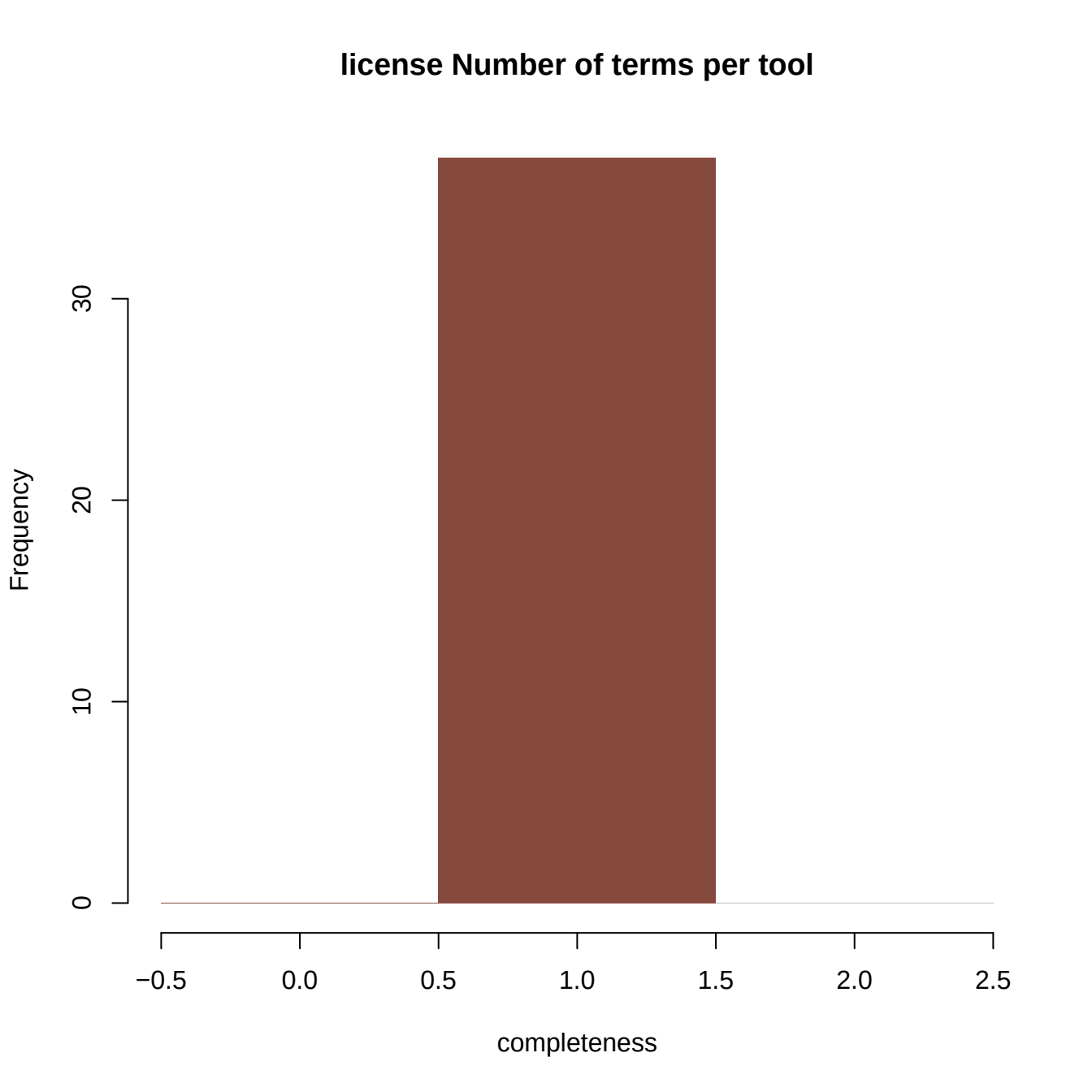

Output format Top 30 terms
8
6
4
2
0
JSON
HTML
TSV
FASTA
PDF
CSV
png
RDF format

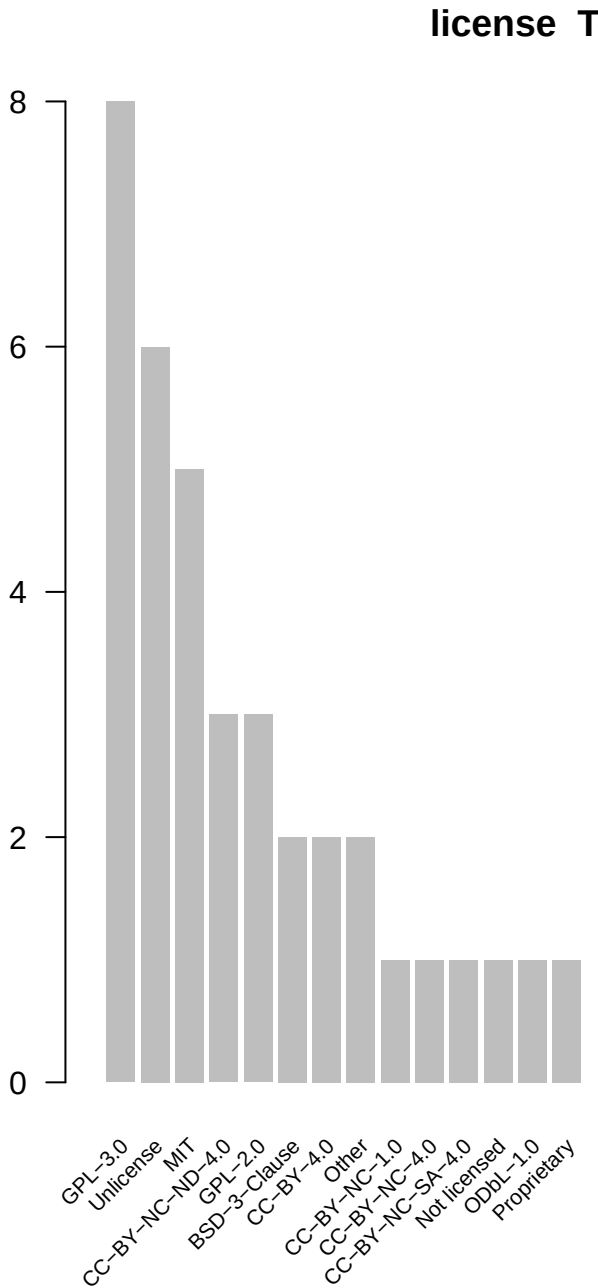

Input data Number of terms per tool
Frequency
0
2
4
6
8
10
12
0
20
40
60
80
completeness

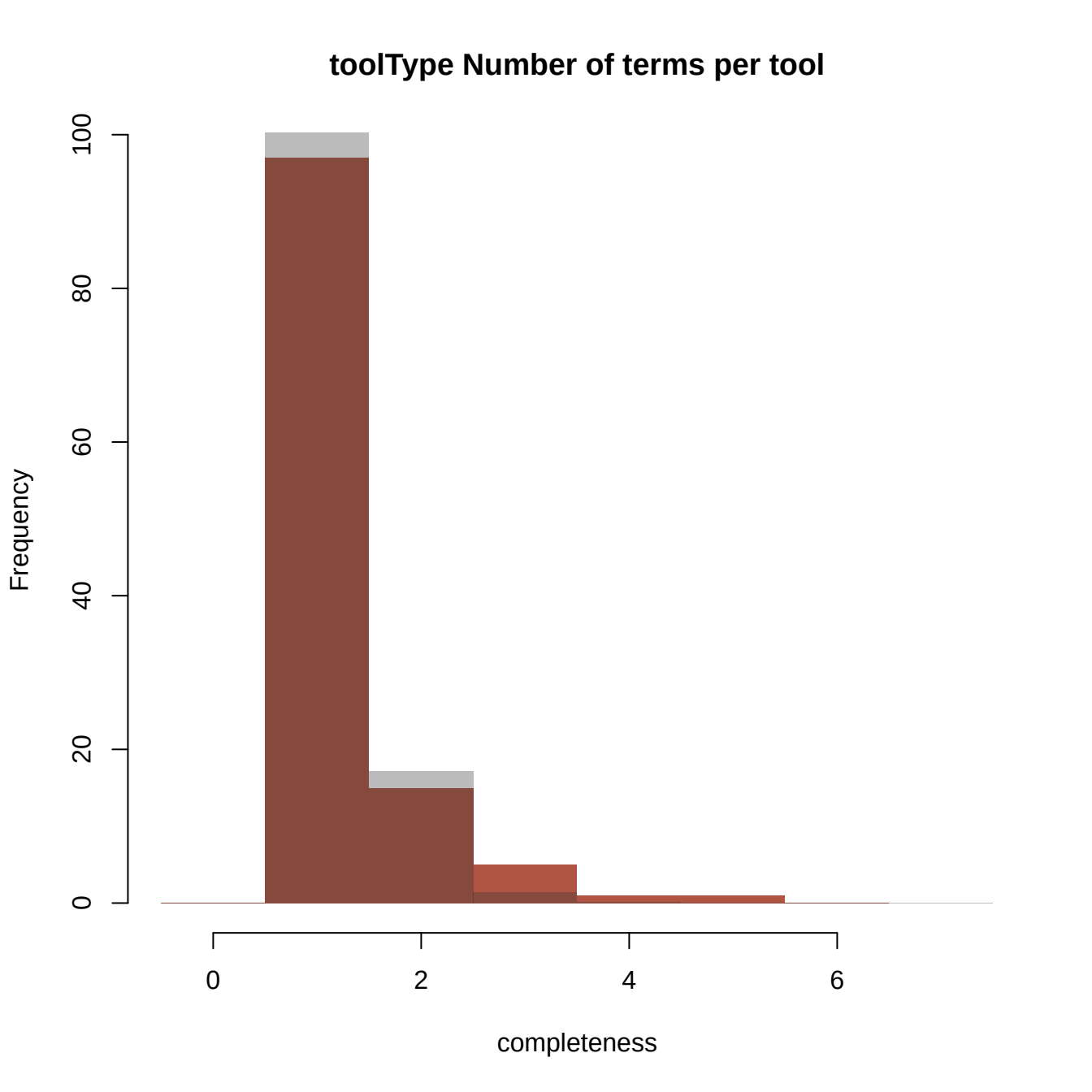

Input data Top 30 terms
5
4
3
2
1
0
Protein sequence
equence record
niProt accession
UniProt keyword
dbSNP ID
Gene ID (NCBI)
equence record
tein ID (DisProt)
Protein name
sequence (raw)
Sequence
uence variations

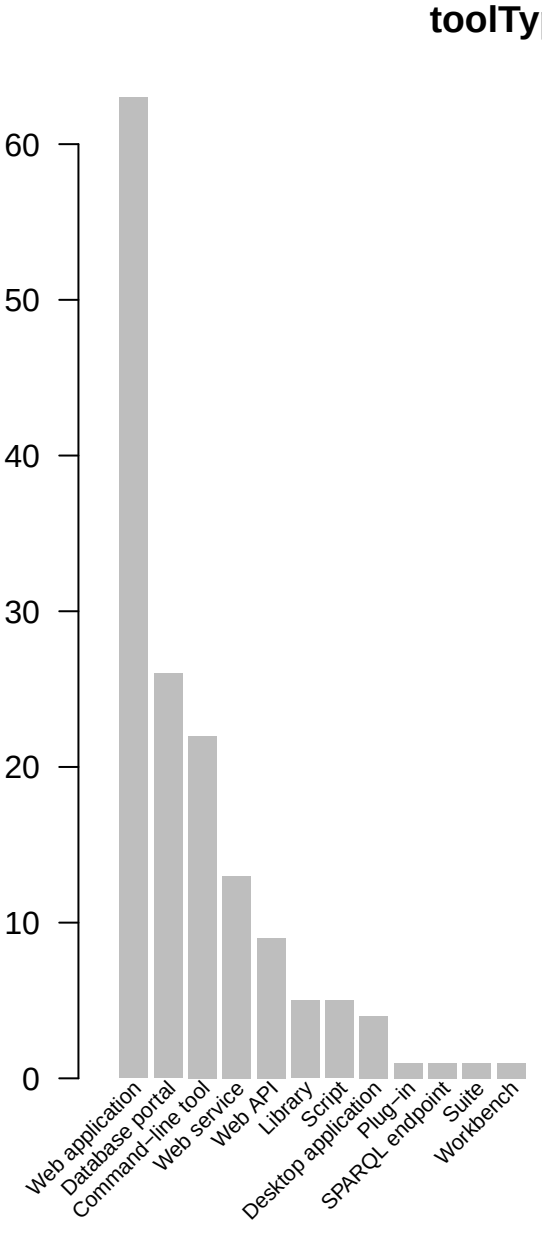

Output data Number of terms per tool
Frequency
0
2
4
6
8
10
12
14
0
10
20
30
40
50
completeness

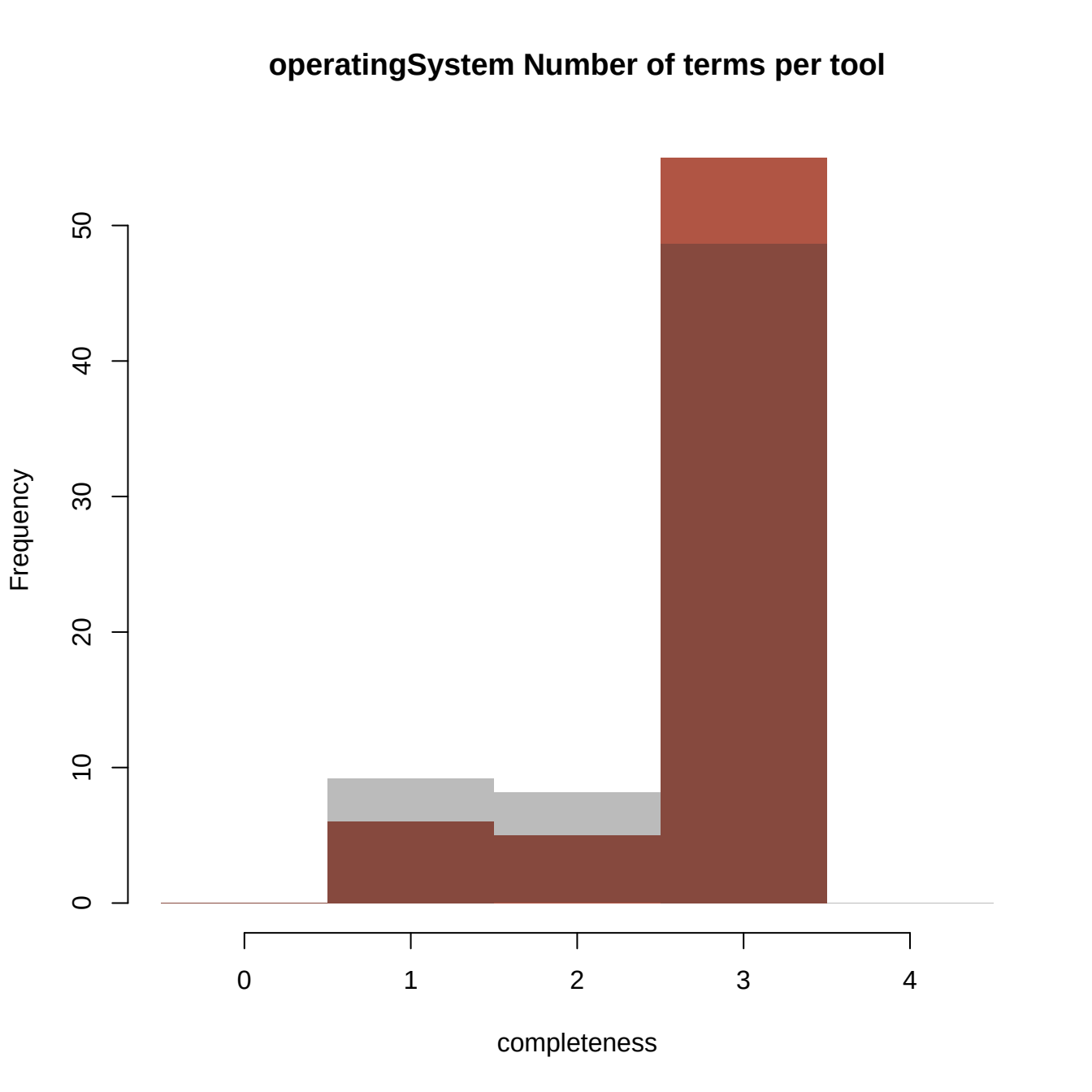

Output data Top 30 terms
7
6
5
4
3
2
1
0
Protein features
Protein report
Sequence features
abase search results
Feature table
file–profile alignment
Protein solubility
Protein structure
Protein structure report
Sequence alignment

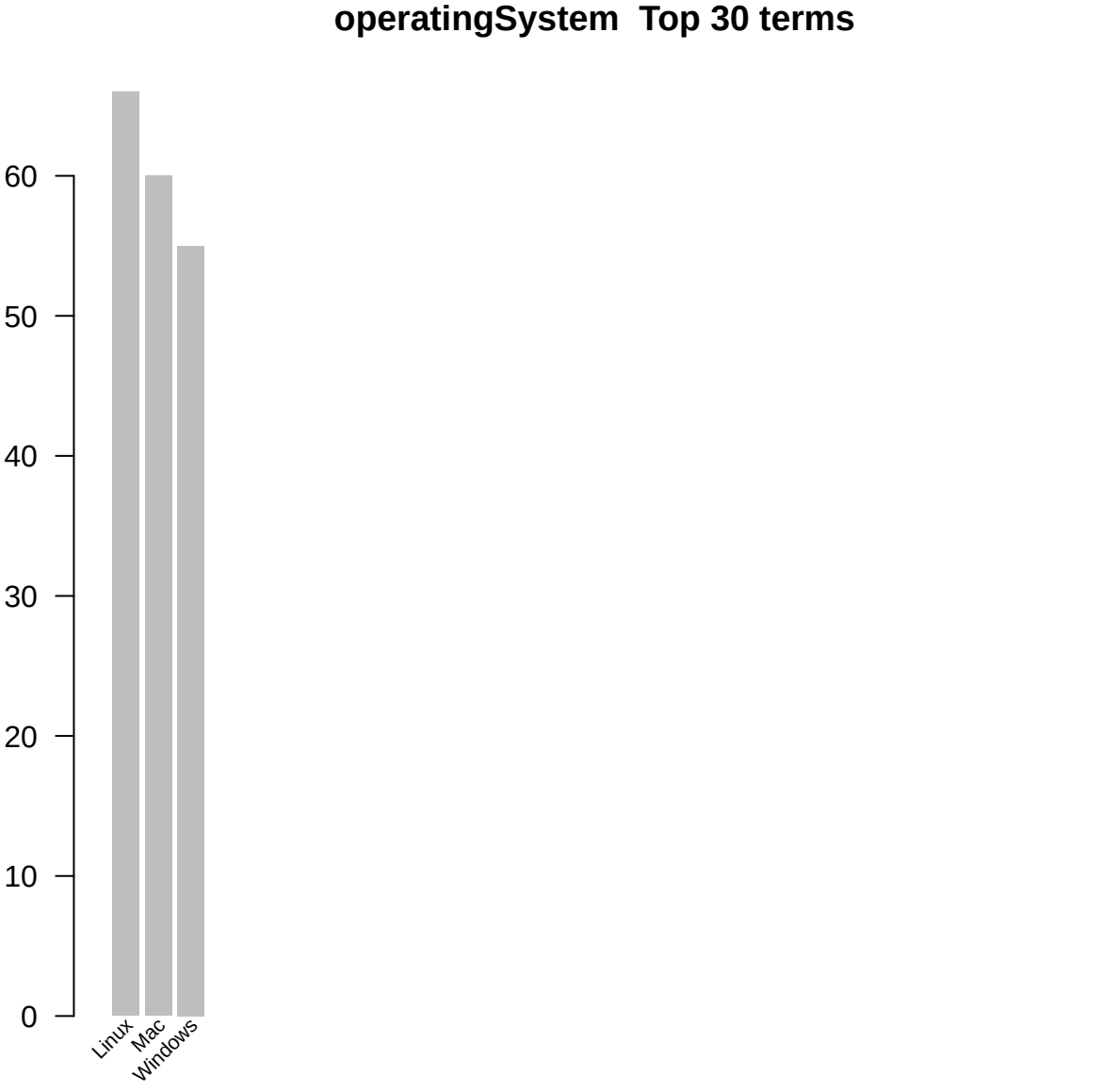

license Number of terms per tool
Frequency
0
10
20
30
−0.5
0.0
0.5
1.0
1.5
2.0
2.5
completeness

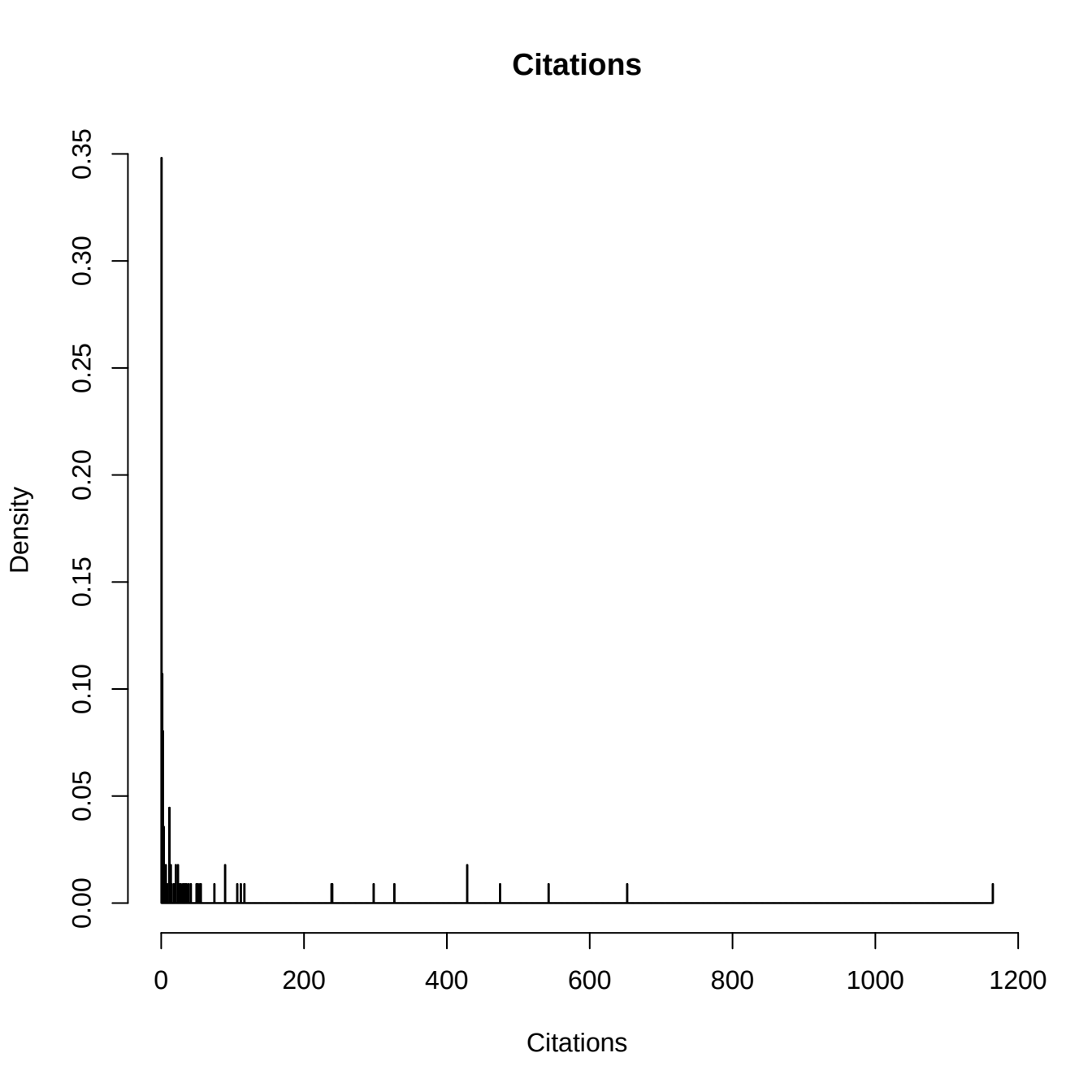

license Top 30 terms
8
6
4
2
0
GPL−3.0
Unlicense
MIT
CC−BY−NC−ND−4.0
GPL−2.0
BSD−3−Clause
CC−BY−4.0
Other
CC−BY−NC−1.0
CC−BY−NC−4.0
CC−BY−NC−SA−4.0
Not licensed
ODbL−1.0
Proprietary

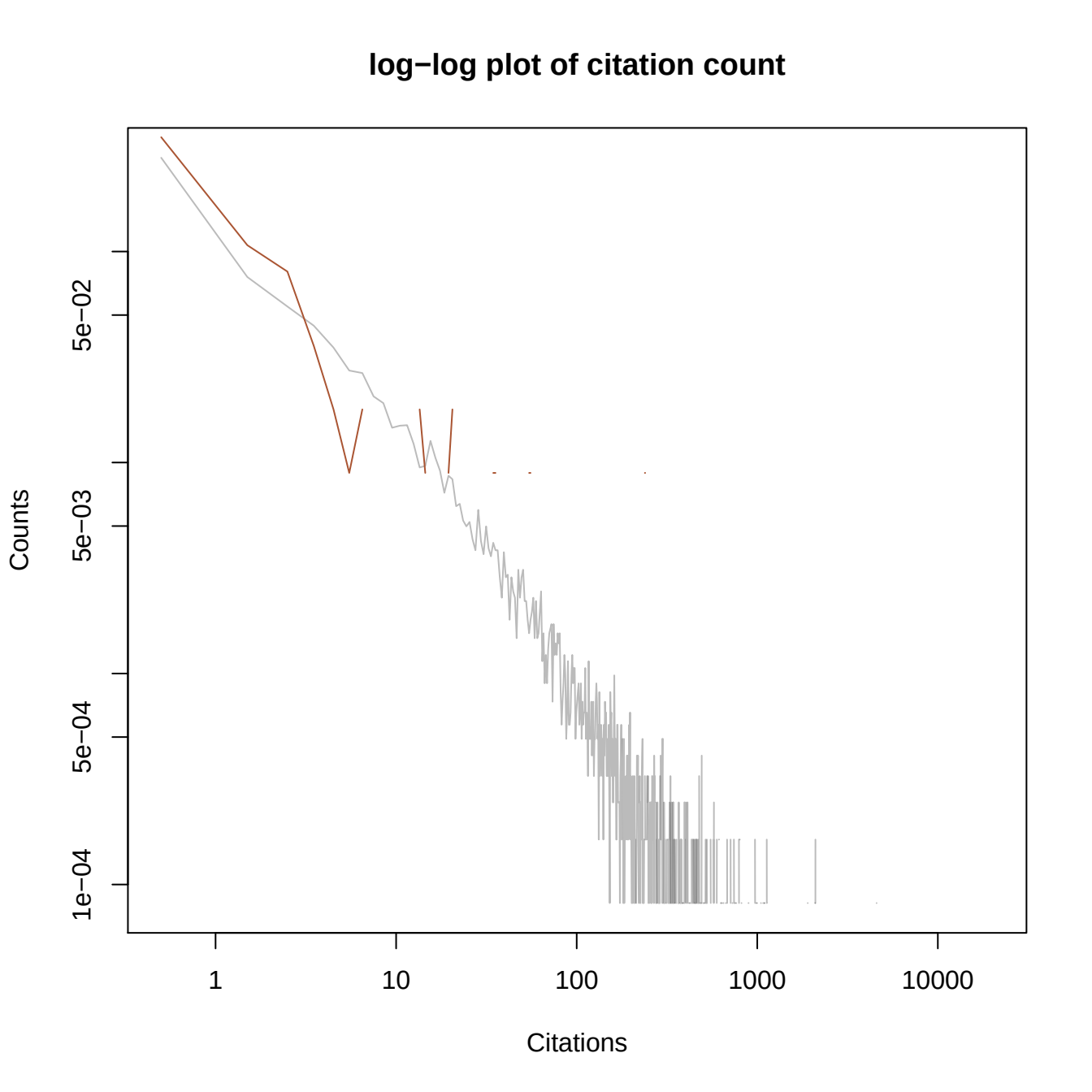

toolType Number of terms per tool
Frequency
0
20
40
60
80
100
0
2
4
6
completeness

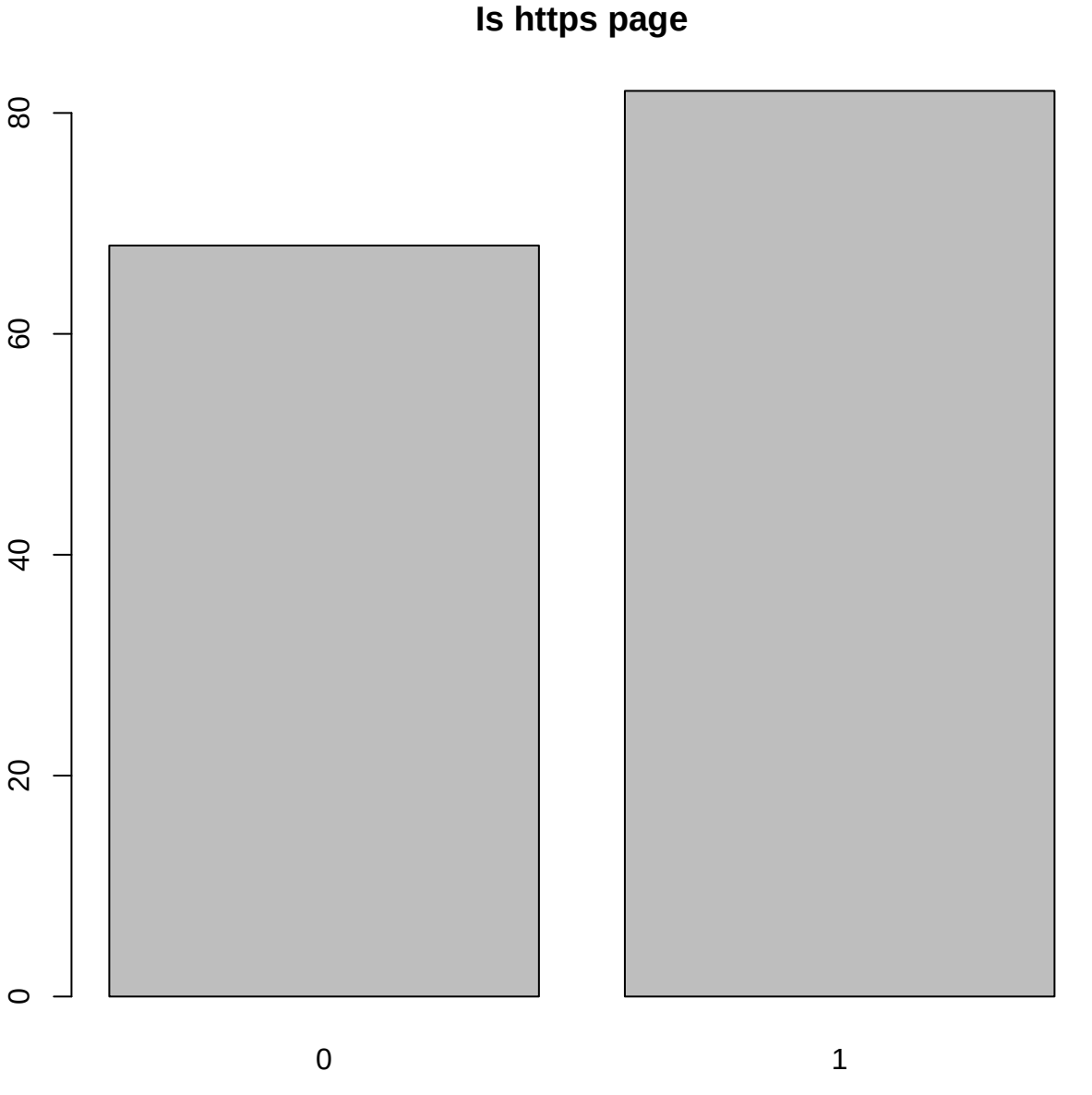

toolType Top 30 terms
60
50
40
30
20
10
0
Web application
Database portal
Command−line tool
Web service
Web API
Library
Script
Desktop application
Plug−in
SPARQL endpoint
Suite
Workbench

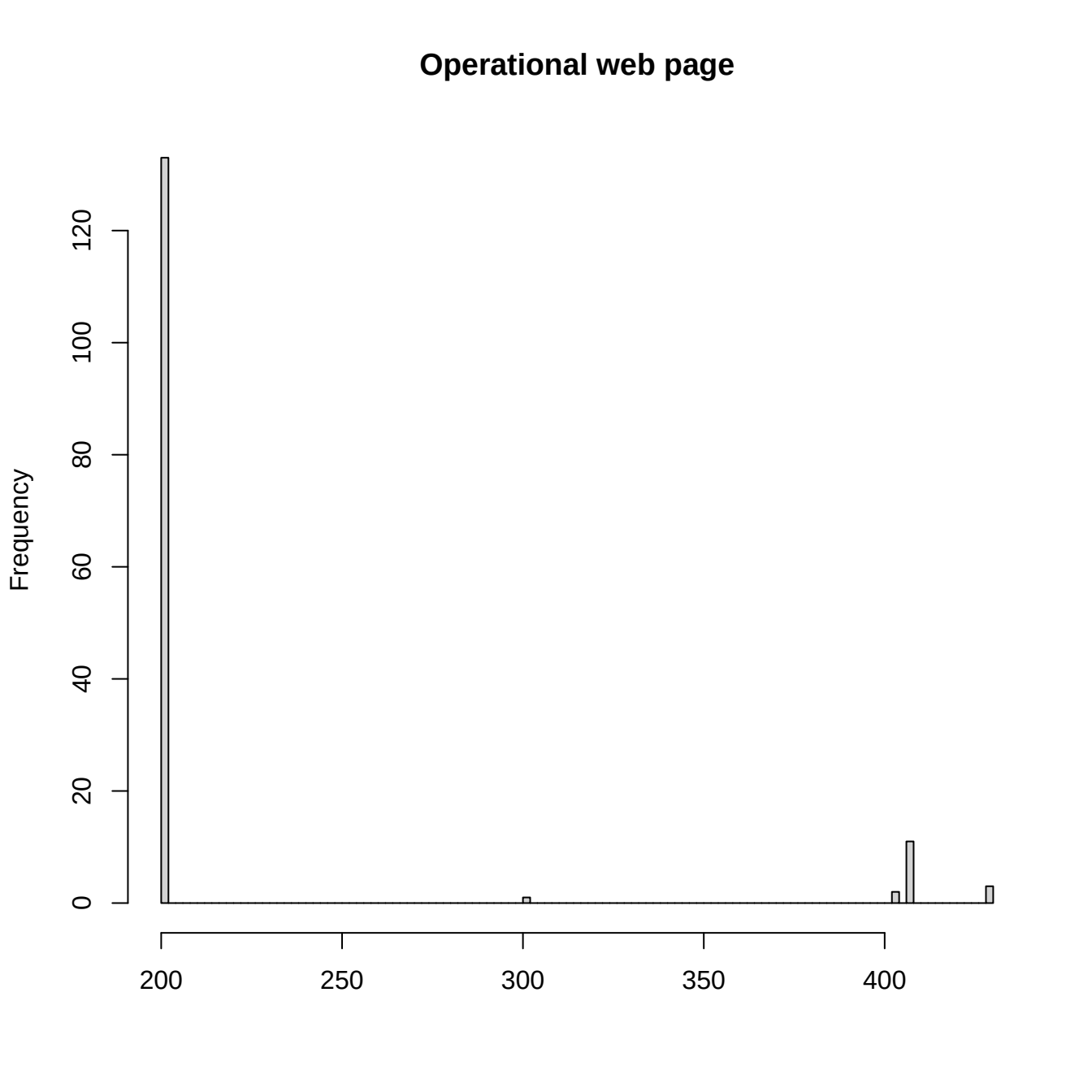

operatingSystem Number of terms per tool
Frequency
0
10
20
30
40
50
0
1
2
3
4
completeness

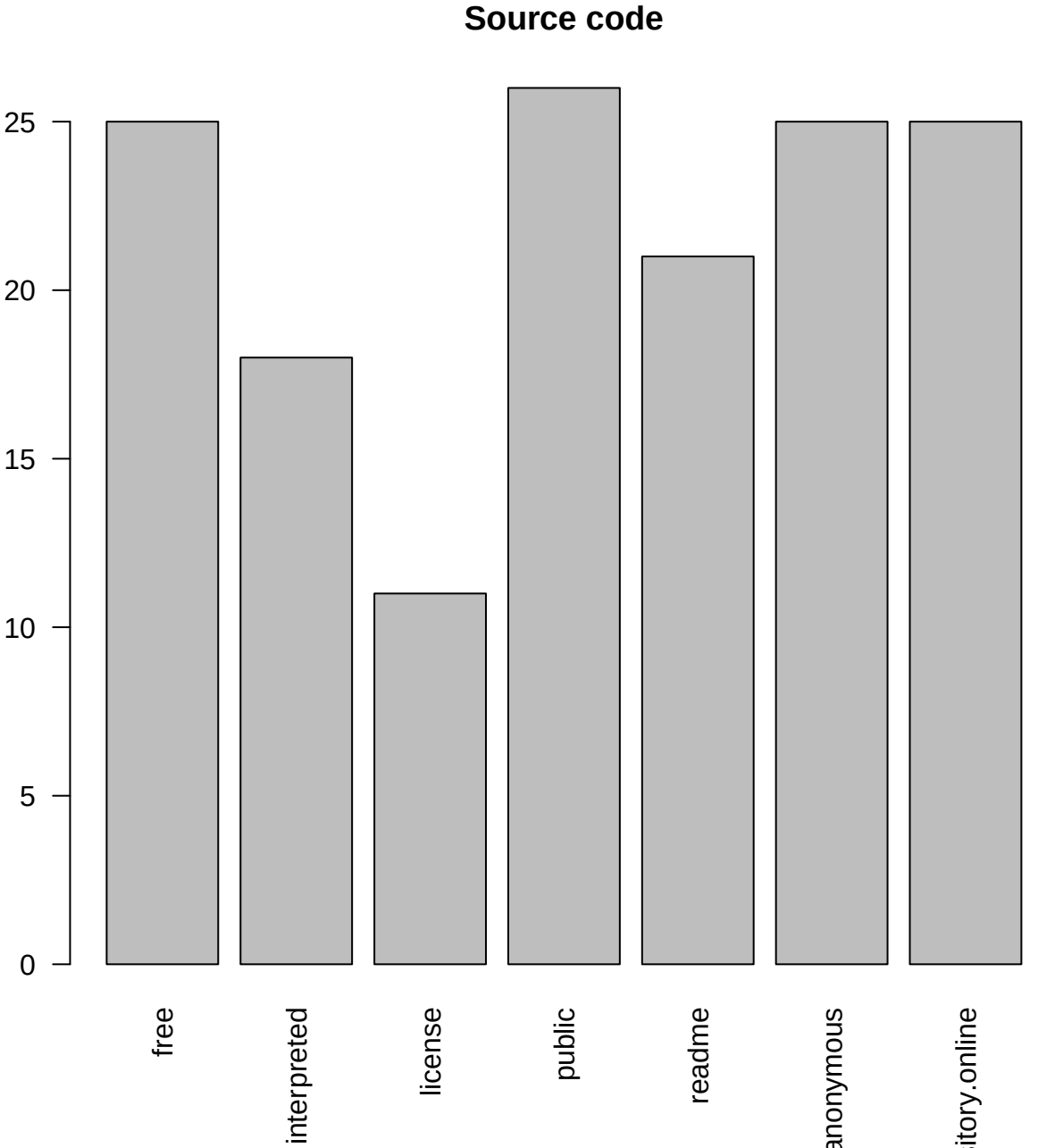

operatingSystem Top 30 terms
60
50
40
30
20
10
0
Linux
Mac
Windows

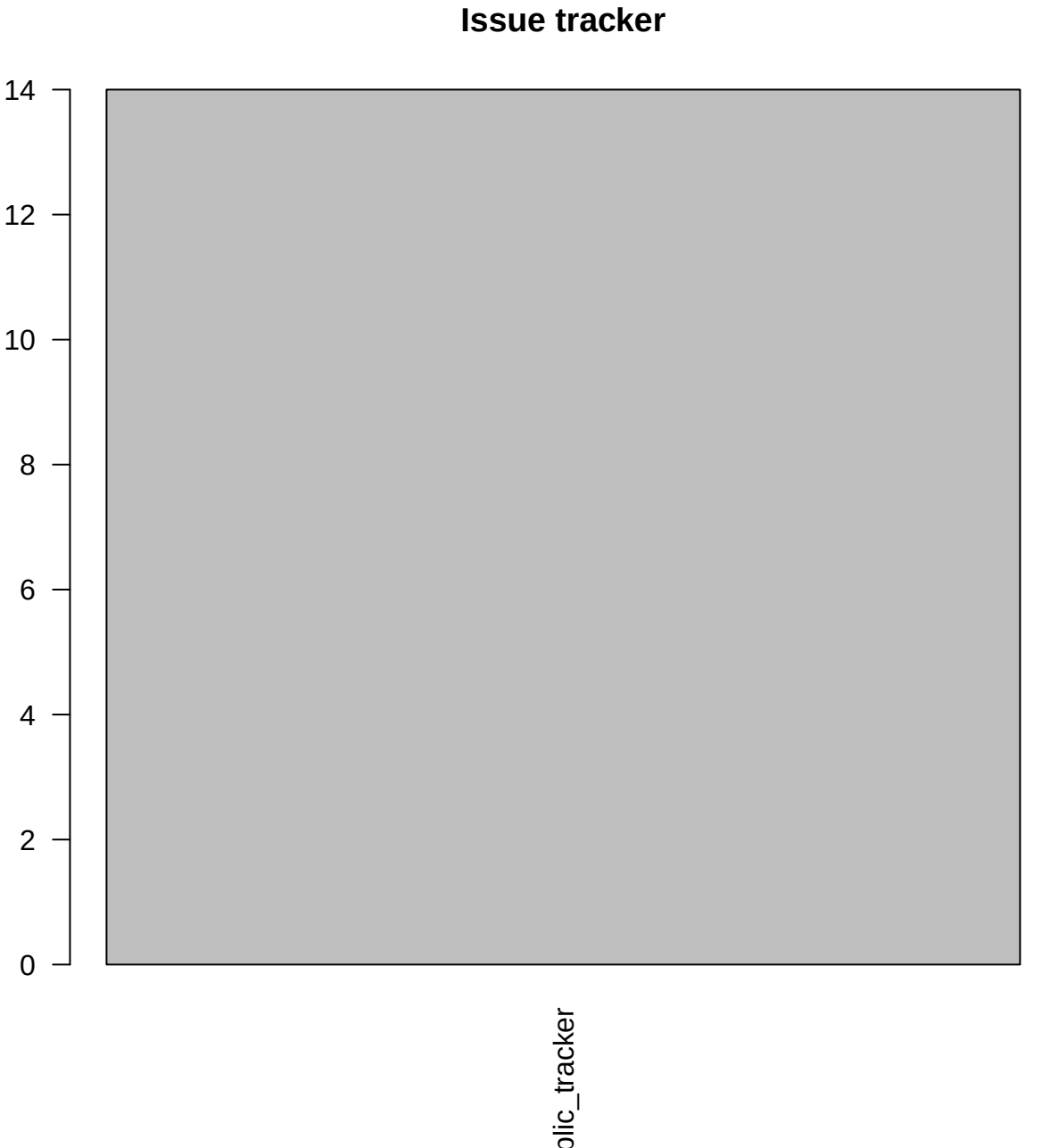

Citations
Density
0.00
0.05
0.10
0.15
0.20
0.25
0.30
0.35
0
200
400
600
800
1000
1200
Citations


log−log plot of citation count
Counts
5e−02
5e−03
5e−04
1e−04
1
10
100
1000
10000
Citations


Is https page
0
20
40
60
80
0
1


Operational web page
Frequency
0
20
40
60
80
100
120
200
250
300
350
400


Source code
25
20
15
10
5
0
free
interpreted
license
public
readme
anonymous
itory.online


Issue tracker
14
12
10
8
6
4
2
0
olic_tracker