

Author Checklist

Genome sequences. Manuscripts should provide:

- A rationale or significance for the sequencing.
- The provenance for the organism sequenced.
- Taxonomic identification down to genus for prokaryotic isolates.
- A description of how the organism was isolated and growth conditions for cultivation. For single-cell amplified genomes, authors should instead supply information about how the cell was identified and isolated.
- Detailed methods for DNA isolation, library preparation, and sequencing (including the technology and chemistry used).
- A description of how the reads were quality controlled.
- Details on how the genome was assembled and, if applicable, annotated.
- A citation and version number for every piece of software used.
- Relevant statistics for the sequencing run (e.g., read length and number of reads in total).
- Relevant statistics for the assembly (e.g., number of contigs and N_{50} values).
- Genome GC content and total size.
- Accession numbers for both the assembly and raw reads that link to publicly available data.

Metagenomics and metatranscriptomic data sets. Manuscripts should provide:

- A rationale or significance for the sequencing.
- The provenance for the sample sequenced and information about the environment.
- Detailed methods for DNA or RNA isolation, library preparation, and sequencing (including the technology and chemistry used).
- A description of how the reads were quality controlled.
- Details on how the metagenome or metatranscriptome was assembled and, if applicable, annotated.
- A citation and version number for every piece of software used.
- Relevant statistics for the sequencing run (e.g., read length and number of reads in total).
- Relevant statistics for the assembly (e.g., number of contigs and N_{50} values).
- Details on how the reads or contigs were binned.
- Accession numbers for both the assembly and raw reads that link to publicly available data.

Amplicon sequence collections. Manuscripts should provide:

- A rationale or significance for the sequencing.
- The provenance for the sample sequenced and information about the environment.
- Detailed methods for DNA or RNA isolation, library preparation, and sequencing (including the technology and chemistry used).
- Relevant statistics for the sequencing run (e.g., read length and number of reads in total).
- A citation and version number for every piece of software used.
- An accession number for the raw reads that links to publicly available data.

Metabolomics data sets. Manuscripts should provide:

- A rationale or significance for the sampling, including experimental factors and study design.
- Details as to how the samples were collected and extracted.
- Details as to the chemistry used to identify metabolites.
- A citation and version number for every piece of software used.
- An accession number that leads to a full data set in a public repository and links to publicly available data.

Proteome data sets. Manuscripts should provide:

- A rationale or significance for the sampling, including experimental factors and study design.
- Details on how the samples were collected, digested, and processed.
- Details as to the peptide identification process.
- A citation and version number for every piece of software used.
- An accession number that leads to a full data set in a public repository and links to publicly available data.

Culture collections, mutant libraries, plasmids, and other genetic resources. Manuscripts should provide:

- A rationale or motivation for the resource and its significance to the research community.
- Details on how the resource was generated.
- A statement on how the resource can be accessed. It is ideal if the resource is placed in a public repository, but if not, the authors must make the resource available to the community.

Databases and software. Manuscripts should provide:

- A description of software implementation and its utility.
- A link to the software as an executable file or a repository.
- A statement on how the resource can be accessed. It is ideal if the resource is placed in a public repository, but if not, the authors must state that they will make the resource available to the community for 10 years after publication.