

Package: XCNV (via r-universe)

July 23, 2026

Type Package

Title Copy Number Variant Annotation and Transparent Scoring

Version 0.0.0.9000

Description Provides bedtools-free copy-number-variant annotation, a portable evaluator and bundled runtime-neutral form of the published X-CNV model of Zhang et al. (2021), and independently authored, provenance-first scoring under the ACMG and ClinGen 2019 constitutional copy-number-variant standard. Large scientific annotation resources remain outside the package payload.

License GPL (>= 2)

URL <https://rgenicsetl.github.io/XCNV/>,
<https://github.com/RGenomicsETL/XCNV>,
<https://github.com/kbvstmd/XCNV>

BugReports <https://github.com/RGenomicsETL/XCNV/issues>

Depends R (>= 4.1.0)

Imports data.table, DBI, duckdb

Suggests roxygen2, tinytest, knitr, rmarkdown

Encoding UTF-8

Roxygen list(markdown = TRUE)

RoxygenNote 8.0.0

LazyData true

LazyDataCompression xz

Config/Needs/website pkgdown

Config/pak/sysreqs xz-utils

Repository <https://rgenicsetl.r-universe.dev>

Date/Publication 2026-07-23 17:19:03 UTC

RemoteUrl <https://github.com/RGenomicsETL/XCNV>

RemoteRef HEAD

RemoteSha dcd13ff68a3d09be0ed7622b4195a20116de18fd

Contents

annotate_cnv	2
classify_cnv_score	3
cnv_2019_criteria	3
cnv_resource_sources	4
derive_cnv_2019_evidence	4
load_xcnv_model	5
predict_cnv	6
read_cnv	6
run_xcnv	7
score_cnv_2019	8
validate_cnv	8
validate_xcnv_resources	9
write_xcnv	9
xcnv_bundled_model	10
xcnv_fixture_resources	10
xcnv_portable_model	11
xcnv_real_example	11
xcnv_resource_manifest	12
xcnv_resources	12
Index	13

annotate_cnv	<i>Annotate CNV records with X-CNV features</i>
--------------	---

Description

Compute the coding, genome-wide, regulatory, haploinsufficiency, and population-frequency features used by X-CNV. Overlap is computed in-process using BED half-open interval arithmetic; the reciprocal 0.5 thresholds used by the upstream frequency step are retained. CNV lengths use the upstream inclusive end - start + 1 convention.

Usage

```
annotate_cnv(x, resources, overlap_backend = c("duckdb", "reference"))
```

Arguments

x	A path, data frame, or matrix containing CNV records.
resources	An xcnv_resources object or an explicit resource directory.
overlap_backend	"duckdb" for the production inequality-join path, or "reference" for the small brute-force correctness oracle.

Value

A data frame in legacy feature order, with the input columns first, the model and auxiliary annotation columns next, and Type.1 and Length at the end.

classify_cnv_score	<i>Convert an ACMG/ClinGen 2019 CNV score to a five-tier classification</i>
--------------------	---

Description

Convert an ACMG/ClinGen 2019 CNV score to a five-tier classification

Usage

```
classify_cnv_score(score)
```

Arguments

score	Numeric score vector.
-------	-----------------------

Value

An ordered factor with the five constitutional CNV classes.

cnv_2019_criteria	<i>ACMG/ClinGen 2019 constitutional CNV criterion contract</i>
-------------------	--

Description

Return the machine-readable criterion identifiers and numeric ranges used by the independently authored scoring lane. This is a scoring contract, not an automatic clinical evidence adjudicator. Dynamic section 5 criteria have an NA default and require an explicit score selected by the curator.

Usage

```
cnv_2019_criteria(cnv_type = c("loss", "gain"))
```

Arguments

cnv_type	One or both of "loss" and "gain".
----------	-----------------------------------

Value

A data frame with allowed per-evidence ranges and aggregate caps.

cnv_resource_sources *Recommended source authorities for constitutional CNV annotation*

Description

Return landing pages and release policies for resources used to construct a provenance-labelled CNV annotation pack. This function does not download data. Staging records the selected release, assembly, and source URL.

Usage

```
cnv_resource_sources()
```

Value

A data frame of source authorities and current pinned releases where the upstream publishes numbered releases.

derive_cnv_2019_evidence
Derive objective ACMG/ClinGen 2019 CNV content evidence

Description

Derive only the criteria determined mechanically by gene content and established ClinGen dosage records: section 1 content, section 2A complete dosage-feature containment, and section 3 protein-coding gene count. Case, phenotype, segregation, population, partial-gene, and inheritance evidence are intentionally not inferred. Inputs are BED half-open relations and must carry source identifiers and releases.

Usage

```
derive_cnv_2019_evidence(  
  cnvs,  
  genes,  
  dosage,  
  functional_biotypes = "protein_coding",  
  established_score = 3  
)
```

Arguments

cnvs	Data frame with cnv_id, chr, start, end, and cnv_type ("loss" or "gain").
genes	Gene relation with chr, start, end, gene_id, biotype, source_id, and source_release.
dosage	ClinGen dosage relation with chr, start, end, record_id, hi_score, ts_score, source_id, and source_release.
functional_biotypes	Biotypes counted as functionally important for sections 1 and 3; defaults to "protein_coding".
established_score	ClinGen score treated as established; defaults to 3.

Value

A provenance-complete evidence data frame accepted by score_cnv_2019().

load_xcnv_model	<i>Load an X-CNV model</i>
-----------------	----------------------------

Description

Load a portable, tabular X-CNV tree model. The portable artifact is exported from the pinned model during resource staging and is evaluated without an XGBoost runtime. Serialized .Rdata, .rds, and native XGBoost files are deliberately rejected; use tools/stage_xcnv_model.R to convert them in an environment where XGBoost is installed.

Usage

```
load_xcnv_model(model = NULL)
```

Arguments

model	A portable model object or path to an xcnv-tree-v1 TSV file. The bundled published model is used when omitted or NULL.
-------	--

Value

A model object suitable for predict_cnv().

predict_cnv	<i>Predict X-CNV MVP scores</i>
-------------	---------------------------------

Description

Annotate CNVs and compute the X-CNV MVP score. With an upstream model this evaluates a staged xcnv-tree-v1 artifact in ordinary R. Installing and running the package does not require XGBoost.

Usage

```
predict_cnv(
  x,
  resources,
  model = NULL,
  overlap_backend = c("duckdb", "reference")
)
```

Arguments

x	A path, data frame, or matrix containing CNV records.
resources	An xcnv_resources object or an explicit resource directory.
model	A model object/path. If NULL, a model in resources is used when present, otherwise the bundled published model is used.
overlap_backend	Passed to annotate_cnv().

Value

A data frame containing the legacy annotation columns and a final MVP_score column.

read_cnv	<i>Read and normalize a CNV table</i>
----------	---------------------------------------

Description

Read a tab-separated BED-like CNV table or normalize an in-memory table. The first four columns are used, matching the published command-line tool; additional columns are ignored.

Usage

```
read_cnv(x)
```

Arguments

x A path, data frame, or matrix containing chromosome, start, end, and CNV type in its first four columns.

Value

A four-column data frame with columns chr, start, end, and type, classed as xcnv_cnv.

run_xcnv	<i>Run X-CNV with a file-compatible interface</i>
----------	---

Description

Read CNVs, compute annotations and MVP scores, and optionally write the output table. Unlike the historical executable, this function never writes beside the input by default and never installs software or downloads data.

Usage

```
run_xcnv(
  input,
  resources,
  model = NULL,
  output = NULL,
  overlap_backend = c("duckdb", "reference")
)

xcnv(
  input,
  resources,
  model = NULL,
  output = NULL,
  overlap_backend = c("duckdb", "reference")
)
```

Arguments

input A CNV path, data frame, or matrix.

resources An xcnv_resources object or explicit resource directory.

model A model object/path, or NULL to use a resource-bundle override when present and otherwise the bundled published model.

output NULL to return results without writing, or one explicit CSV path.

overlap_backend Passed to annotate_cnv().

Value

A prediction data frame; the same object is returned invisibly when output is non-NULL.

score_cnv_2019	<i>Score transparent ACMG/ClinGen 2019 constitutional CNV evidence</i>
----------------	--

Description

The input is an auditable evidence relation. Every row names its CNV, deletion/gain context, criterion, numeric score, source, source release, and evidence record. Repeated case evidence is capped by the published criterion group maximum/minimum. Evidence selection remains explicit and is never inferred from free text or phenotype similarity by this function.

Usage

```
score_cnv_2019(evidence, detail = c("summary", "groups"))
```

Arguments

- evidence Data frame with cnv_id, cnv_type, criterion, score, source_id, source_release, and evidence_id; optional logical selected defaults to TRUE.
- detail "summary" for one row per CNV or "groups" for the capped criterion-group audit relation.

Value

A data frame.

validate_cnv	<i>Validate CNV records</i>
--------------	-----------------------------

Description

Validate the first four columns of a CNV table using the input rules of the pinned X-CNV executable. Chromosome names may use the executable's leading chr convention, and coordinates are retained as integer values.

Usage

```
validate_cnv(x)
```

Arguments

- x A path, data frame, or matrix containing chromosome, start, end, and CNV type in its first four columns.

Value

TRUE invisibly when validation succeeds; otherwise an ordinary R error is raised.

```
validate_xcnv_resources
```

Validate an X-CNV resource bundle

Description

Check that a local directory contains the named X-CNV annotation and/or model resources. Validation never downloads files and never changes the working directory.

Usage

```
validate_xcnv_resources(path = NULL, require = "annotations")
```

Arguments

path	Resource directory. If omitted, XCNV_RESOURCE_DIR must be set.
require	Resource groups to require: "annotations" (the default), "model", or "all". A model file is required only when "model" or "all" is requested; otherwise prediction uses the bundled model.

Value

An object of class xcnv_resources, containing the normalized root and resolved resource paths.

```
write_xcnv
```

Write X-CNV-compatible output

Description

Write a prediction data frame as a comma-separated X-CNV output table. A file is written only when an explicit output path is supplied.

Usage

```
write_xcnv(x, output)
```

Arguments

x	A data frame returned by predict_cnv().
output	Explicit output path.

Value

The normalized output path, invisibly.

xcnv_bundled_model	<i>Return the bundled published X-CNV model</i>
--------------------	---

Description

Return the bundled published X-CNV model

Usage

```
xcnv_bundled_model()
```

Value

A validated xcnv_portable_model object.

xcnv_fixture_resources	<i>Locate the bundled deterministic fixture</i>
------------------------	---

Description

Resolve the small, non-scientific fixture shipped for tests and examples. It is not a substitute for the Zhang et al. annotations or model.

Usage

```
xcnv_fixture_resources()
```

Value

An object of class xcnv_resources.

xcnv_portable_model	<i>Published X-CNV portable model</i>
---------------------	---------------------------------------

Description

The fitted X-CNV classifier from Zhang et al. (2021), converted once from its serialized XGBoost representation to the package's runtime-neutral xcnv-tree-v1 relation.

Format

An xcnv_portable_model list containing the logistic base score, 30 feature names, and a 41-row tree relation.

Source

Zhang et al. (2021), [doi:10.1186/s13073021009454](https://doi.org/10.1186/s13073021009454).

xcnv_real_example	<i>Load the bundled real X-CNV example</i>
-------------------	--

Description

Return ClinVar allele 72461, a pathogenic deletion containing SOX2, with the case-limited GRCh37 annotation relations needed to run X-CNV and the objective ACMG/ClinGen 2019 content-evidence lane. The annotation files are subsets for this case and must not be used for unrelated CNVs.

Usage

```
xcnv_real_example()
```

Value

A list containing cnv, clinvar, resources, genes, and dosage.

xcnv_resource_manifest	<i>Show the X-CNV resource contract</i>
------------------------	---

Description

Return the file names and resource groups understood by the package. Large annotations are external resources; the model entry is an optional override for the bundled classifier.

Usage

```
xcnv_resource_manifest()
```

Value

A data frame describing required resource keys, file names, and groups.

xcnv_resources	<i>Discover an X-CNV resource bundle</i>
----------------	--

Description

Resolve and validate a resource directory. With no path, only the explicit XCNV_RESOURCE_DIR environment variable is consulted; the current working directory is never searched implicitly.

Usage

```
xcnv_resources(path = NULL, require = "annotations")
```

Arguments

- path Resource directory, or NULL to use XCNV_RESOURCE_DIR.
- require Resource groups to require.

Value

An object of class xcnv_resources.

Index

`annotate_cnv`, [2](#)

`classify_cnv_score`, [3](#)
`cnv_2019_criteria`, [3](#)
`cnv_resource_sources`, [4](#)

`derive_cnv_2019_evidence`, [4](#)

`load_xcnv_model`, [5](#)

`predict_cnv`, [6](#)

`read_cnv`, [6](#)
`run_xcnv`, [7](#)

`score_cnv_2019`, [8](#)

`validate_cnv`, [8](#)
`validate_xcnv_resources`, [9](#)

`write_xcnv`, [9](#)

`xcnv (run_xcnv)`, [7](#)
`xcnv_bundled_model`, [10](#)
`xcnv_fixture_resources`, [10](#)
`xcnv_portable_model`, [11](#)
`xcnv_real_example`, [11](#)
`xcnv_resource_manifest`, [12](#)
`xcnv_resources`, [12](#)