

Overview of vignettes for copy number estimation

Rob Scharpf

April 15, 2025

The workflow for copy number analyses in the `crmm` package includes preprocessing and genotyping of the raw intensities followed by estimation of parameters for copy number estimation using `crmmCopynumber`. Supported platforms are those for which a corresponding annotation package is available. Table ?? provides an overview of the available vignettes pertaining to copy number estimation. These vignettes are located in the `inst/scripts` subdirectory of the `crmm` package. HapMap datasets are used to illustrate the workflow and are not provided as part of the `crmm` package. Users wishing to reproduce the analysis should download the HapMap CEL files (Affymetrix) or the `idat` files (Illumina) and modify the paths to the raw data files as appropriate.

Vignette	Platform	Annotation package	Scope
Infrastructure	Affy/Illumina		The <code>CNSet</code> container / large data support using the <code>ff</code> package
AffyGW	Affy 5.0, 6.0	<code>genomewidesnp5crmmml</code> , <code>genomewidesnp6Crlmm</code>	Preprocessing, genotyping, CN estimation
IlluminaPreprocessCN	Illumina	several [†]	Preprocessing, genotyping, CN estimation

Table 1: Vignettes for copy number estimation. [†] See `annotationPackages()` for a complete listing of supported Illumina/Affy platforms

The `Infrastructure` vignette provides additional details on the `CNSet` container used to organize the processed data as well as a brief discussion regarding large data support through the `ff` package.