

GATK gCNV Usage Documentation

GATK CNV Team

January 12, 2018

1 What is GATK gCNV?

GATK gCNV is a read-depth based germline copy-number variant (CNV) calling tool.

2 Overview of GATK gCNV pipeline

The necessity of data modeling in general, ploidy determination, and data preparation in particular for GATK gCNV.

2.1 COHORT mode

todo.

2.2 CASE mode

todo.

2.3 Scalability and task sharding

3 Setting up GATK gCNV runtime environment

todo. setting up the conda environment for gcnvkernel.

4 Preparing the data and resources

todo.

5 Components of GATK gCNV

5.1 DetermineGermlineContigPloidy

todo.

5.1.1 Purpose of the Tool

todo.

5.1.2 Operation Modes

todo.

5.1.3 Description of Arguments

todo.

5.2 GermlineCNVCaller

todo.

5.2.1 Purpose of the Tool

todo.

5.2.2 Operation Modes

todo.

5.2.3 Description of Arguments

todo.

5.3 GermlineCNVPostProcessor

todo.

5.4 Description of Hybrid ADVI Inference Arguments

todo.

6 Running GATK gCNV pipeline manually

todo.

7 Running GATK gCNV via WDL scripts

todo.