

Lightning - Story #14085

[CWL] Create VCF + BED to gVCF pipeline

08/21/2018 12:15 PM - Abram Connelly

Status:	Closed	Start date:	
Priority:	Normal	Due date:	
Assigned To:	Abram Connelly	% Done:	0%
Category:		Estimated time:	0.00 hour
Target version:			

Description

Create a CWL pipeline to convert a VCF file with an associated BED file of homozygous reference regions to a gVCF file. The short term goal of the gVCF is to be fed into a Lightning pipeline for conversion to FastJ and ultimately Compact Genome Format (CGF).

There is preliminary work on doing this conversion with the command line tool [vcfbed2homref](#). From a high level perspective, the CWL pipeline should just be a wrapper around this tool. Practically, there are some issues to work out, such as whether the input VCF is formatted properly (one large VCF file or multiple VCF files split by chromosome?) and making sure the resulting VCF file is indexed properly.

Part of the resolution of this ticket will be understanding what the expectations on input VCF + BED files are as well as understanding what the output requirements of the resulting gVCF files are.

History

#1 - 08/21/2018 01:07 PM - Abram Connelly

- Status changed from New to Closed

I completely forgot this had already been done. See [#13425](#).

- [CWL VCF + BED to gVCF](#)