============================= Methods Summary ===================================

Read data was assembled using ALLPATHS-LG revision R48559. Pilon version 1.13 was run to correct base errors and fill gaps. Contaminant adapter and duplicated scaffolds were identified and removed. GAEMR was run to produce assembly chart and metric tables that were examined during the manual analysis process.  Manual analysis included the removal of contigs from off-target genomes, primary *Wolbochia* and *Mycoplasma*.

============================= URLs and Citations ===================================

ALLPATHS-LG manual, download, and citation:

<ftp://ftp.broadinstitute.org/pub/crd/ALLPATHS/Release-LG/AllPaths-LG_Manual.pdf>

<ftp://ftp.broadinstitute.org/pub/crd/ALLPATHS/Release-LG/latest_source_code/LATEST_VERSION.tar.gz>

Gnerre S, MacCallum I, Przybylski D, Ribeiro F, Burton J, Walker B, Sharpe T, Hall G, Shea T, Sykes S, Berlin A, Aird D, Costello M, Daza R, Williams L, Nicol R, Gnirke A, Nusbaum C, Lander ES, Jaffe DB. 2010. [High-quality draft assemblies of mammalian genomes from massively parallel sequence data.](http://www.pnas.org/content/early/2010/12/20/1017351108.abstract) Proc Natl Acad Sci U S A. 2011 Jan 25;108(4):1513-8.

GAEMR download:

<http://www.broadinstitute.org/software/gaemr/>

Pilon download and citation:

<https://github.com/broadinstitute/pilon/releases>

Bruce J. Walker, Thomas Abeel, Terrance Shea, Margaret Priest, Amr Abouelliel, Sharadha Sakthikumar, Christina A. Cuomo, Qiandong Zeng, Jennifer Wortman, Sarah K. Young, Ashlee M. Earl (2014) Pilon: An Integrated Tool for Comprehensive Microbial Variant Detection and Genome Assembly Improvement. PLoS ONE 9(11): e112963. doi:10.1371/journal.pone.0112963

======================== ALLPATHS-LG directory setup ===============================

Create UNIX directories according to required ALLPATHS structure, /PRE/REFERENCE/DATA\_SUBDIR/RUN/ASSEMBLIES/SUBDIR (see page 7 of the manual). In the example path here:

/cil/shed/apps/internal/ale/production/projects/X966/ SM-AJDMW/ SM-AJDMW\_allpaths\_39f30j\_59742/run/ASSEMBLIES/test/

the breakdown of PRE, REFERENCE, etc.. is as follows:

PRE= /cil/shed/apps/internal/ale/production/projects/X966

REFERENCE=SM-AJDMW

DATA\_SUBDIR:=SM-AJDMW\_allpaths\_39f30j\_59742

RUN=run

ASSEMBLIES= ASSEMBLIES

SUBDIR= test

Within the REFERENCE directory create a “cache” directory. In the example above this would be “/cil/shed/apps/internal/ale/production/projects/X966/SM-AJDMW/cache”.

Within the cache directory create file named “in\_libs.csv” with following contents:

library\_name, library\_id, project\_name, organism\_name, type, paired, frag\_size, frag\_stddev, insert\_size, insert\_stddev, read\_orientation, genomic\_start, genomic\_end, import\_date

SM-AJDMW.jumps, 0001, SM-AJDMW, Ant sample SM-AJDMW, jumping, 1, , , 3200, 800, outward, 0, 0, 2017-06-29 09:11:51

SM-AJDMW.frags, 0002, SM-AJDMW, Ant sample SM-AJDMW, fragment, 1, 180, 20, , , inward, 0, 0, 2017-06-29 09:11:52

Within cache directory create file named “in\_groups.csv” with following contents:

group\_name, library\_name, file\_name

H7N27ADXY.2.SM-AJDMW.jumps, SM-AJDMW.jumps, /btl/data/walkup/Ant/data/H7N27ADXY/C1-210\_2016-08-09\_2016-08-10/2/SM\_AJDMW/H7N27ADXY.2.unmapped.bam

H3NHTADXY.1.SM-AJDMW.jumps, SM-AJDMW.jumps, /btl/data/walkup/Ant/data/H3NHTADXY/C1-210\_2016-07-20\_2016-07-21/1/SM\_AJDMW/H3NHTADXY.1.unmapped.bam

H3NHTADXY.2.SM-AJDMW.frags, SM-AJDMW.frags, /btl/data/walkup/Ant/data/H3NHTADXY/C1-210\_2016-07-20\_2016-07-21/2/SM\_AJDMW/H3NHTADXY.2.unmapped.bam

HVGCWBCXX.1.SM-AJDMW.frags, SM-AJDMW.frags, /btl/data/walkup/Ant/data/HVGCWBCXX/SM\_AJDMW/HVGCWBCXX.1.unmapped.bam

H7N27ADXY.1.SM-AJDMW.frags, SM-AJDMW.frags, /btl/data/walkup/Ant/data/H7N27ADXY/C1-210\_2016-08-09\_2016-08-10//1/SM\_AJDMW/H7N27ADXY.1.unmapped.bam

H3VHLADXY.1.SM-AJDMW.frags, SM-AJDMW.frags, /btl/data/walkup/Ant/data/H3VHLADXY/C1-218\_2016-07-08\_2016-07-11//1/SM\_AJDMW/H3VHLADXY.1.unmapped.bam

HVNY5BCXX.1.SM-AJDMW.jumps, SM-AJDMW.jumps, /btl/data/walkup/Ant/data/HVNY5BCXX/C1-210\_2016-09-22\_2016-09-27/1/SM\_AJDMW/HVNY5BCXX.1.unmapped.bam

HVNY5BCXX.2.SM-AJDMW.frags, SM-AJDMW.frags, /btl/data/walkup/Ant/data/HVNY5BCXX/C1-210\_2016-09-22\_2016-09-27/2/SM\_AJDMW/HVNY5BCXX.2.unmapped.bam

Note that in the above groups.csv file one would need to give the current path to BAM files.

Once the above directories and files are created the user may then proceed to the steps outline in the attached excel file.