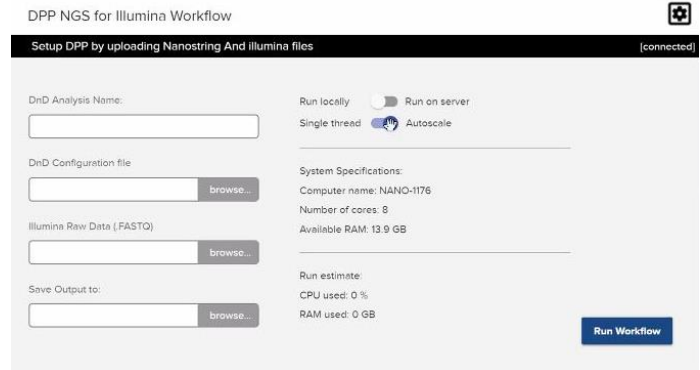


# GeoMx NGS Pipeline v1.0 Software Release Notes

The **GeoMx NGS Pipeline software v1.0** (also known as the DND software) was developed by NanoString and released in July 2020. The Pipeline uses a series of open-source components to process RNA- sequencing files (FASTQ files) from Illumina sequencers and produce *.dcc* files, which can then be uploaded to the GeoMx DSP system for data analysis. The following is a summary of v1.0 features and known issues. Contact [GeoMxSupport@nanoString.com](mailto:GeoMxSupport@nanoString.com) for additional information and assistance.



## Features

- Processes FASTQ files from Illumina sequencers according to parameters defined in the Configuration File (which is generated from the GeoMx DSP run).
- Outputs *.dcc* files, which can then be uploaded to the GeoMx DSP system for data analysis.
- A graphical user interface (GUI) which can be installed on Windows or MacOS. This GUI allows the user to browse to input files needed, designate an output location, choose server or local processing, and monitor progress.
- Users may also choose to use command line interface (CLI) to run the GeoMx NGS Pipeline.
- The Pipeline program itself may run locally (if running 12 FASTQ files or fewer) or on a server (most common). Linux and Amazon Web Services (AWS) servers are supported.
- The Pipeline program can be run on AWS with AWS Batch configuration which allows highly parallelized batch processing for the fastest processing of large data sets with many samples.

## Known Issues

- Some errors result in empty *.dcc* files. Check *.dcc* file sizes and the summary.txt file to ensure files were processed as expected, especially for AWS batch runs as errors from separate batch jobs are not always consolidated into logs.
- Ensure that your FASTQ files do not contain underscore characters () in the sequence identifier lines, as this may cause processing to fail, stall, or produce count files with errors. You may have

**FOR RESEARCH USE ONLY. Not for use in diagnostic procedures.**

© 2020 NanoString Technologies, Inc. All rights reserved.

NanoString, NanoString Technologies, the NanoString logo, nCounter and GeoMx are trademarks or registered trademarks of NanoString Technologies, Inc., in the United States and/or other countries. All other trademarks and/or service marks not owned by NanoString that appear in this document are the property of their respective owners.

underscores in your file names if your Illumina instrument name contains an underscore; if this is the case, modify your instrument name to remove any underscores. In addition, bcl to FASTQ file conversion software may be set to use custom headers; ensure the software you use to generate FASTQ files does not add any underscores to the sequence identifier lines.

- Do not modify FASTQ filenames from the [Illumina FASTQ file naming conventions](#). If possible, retain the naming from the sample sheets exported from DSP, which incorporate the plate barcode. If you must change the sample names, ensure you do not add extra underscores, as the GeoMx NGS Pipeline software uses underscores to parse the Illumina-written portion of the file names to identify well/sample, lane, and read. In addition, if the sample name portions of your filenames are modified from the default pipeline workflow, you will need to update the names in your config file accordingly before you begin your NGS processing pipeline run.
- Submitting jobs from the GUI to a server may result in errors if there is already a run processing on the server.
- If logs from AWS batch run are not present, you may be able to view them on AWS cloudwatch. See documentation from AWS for more information.
- The GeoMx NGS Pipeline running on the server only listens on port 5000 and there is no argument to change this. This may cause a problem if there are other apps on server or network limiting the availability of that port.
- While the GeoMx NGS Pipeline is processing, it writes many large intermediate files to disc. Therefore, you will need up to 10x more space than your initial file set size on the computer where processing is occurring to complete processing. These intermediate files are removed automatically unless you modify the config file to “keepStepFile=True”.
- If inspecting your *.dcc* files in a text editor, you may notice lines for “tamperedIni,No” and “SOMEHASH100000000000”. These are not currently functional and are placeholders for a future feature; they do not affect your data.
- Processing FASTQ files for only a subset of the ROIs listed in the GeoMx NGS Pipeline (DND) configuration file will result in empty *.dcc* files for the missing ROIs. Uploading these empty *.dcc* files will overwrite any data that may exist for these segments from a previous DND run. The GeoMx NGS Pipeline configuration file can be revised to exclude any ROIs for which FASTQ files were previously processed. The resulting *.dcc* zip uploaded should then only update segments associated with blank data.
- When running on AWS batch, some jobs may silently fail, returning *.dcc* files with zero aligned and deduplicated reads or not returning some *.dcc* files at all, even though the AWS job console indicates that the job succeeded. Always review the summary.txt file to note any unexpected errors or zero reads for any intermediate steps for ROIs. Logs for a given job ID may be retrieved from AWS cloudwatch in order to troubleshoot.
- If executing a run from the GUI, any errors or failures will be indicated in the notification menu and listed by ROI name.