

NBDPS Sequencing Study External Quality Assessment Protocol
Approved May 2013

Create a standard protocol using control samples for external quality assessment (EQA). The primary objective of establishing EQA is to ensure that each lab actively involved in testing NBDPS samples is proficient in their respective laboratory techniques independent of the source material or extraction procedure.

EQA Samples: Pre-Characterized (from NIGMS or NHGRI Repositories at Coriell)

- Composition of Pre-Characterized Samples:
 - Purchase sample(s) that have been deeply sequenced by the 1000 Genomes consortium with publicly available data.
 - Coriell ships to each investigator 1 sample. Investigators are blinded to sample and will sequence the sample using the same methods/platforms that will be used with NBDPS samples. Results will be reported back to CDC and compared to results from other labs and the published results.

- Use of Pre-Characterized Samples:
 - Pre-characterized samples will be used prior to initiating NBDPS sequencing projects and will be completed one time per lab per project. For targeted sequencing projects, laboratories should report all sequencing results. For exome or full genome sequencing projects, laboratories should report results from a region agreed upon by the GAWG. Results should be aligned to the reference genome and submitted in BAM or SAM file format. Results will be compared between labs and with published results.

- Pre-Characterized Samples Will Allow:
 - Comparison to published third-party results
 - Inter-lab comparison of sequencing results when possible

Standards Required to Pass EQA:

- 90% concordance between successful sequencing data from regions previously sequenced at $\geq 10X$ depth for:
 - inter-lab sequencing data
 - pre-characterized DNA and published third-party results

- If inter-lab results for assays performed in common are discordant, results from assays performed on pre-characterized samples will be compared to third party published results to determine if a lab needs to identify and resolve problems.

- If a laboratory does not pass EQA standards, they must discontinue all sequencing assays and repeat EQA. If the lab does not pass EQA standards a second time, no manuscripts will be completed until the problems are identified and resolved.

Additional Items:

- Results reported to CDC are final (e.g., if errors are made transcribing data to results template and the data do not meet the standards required to pass EQA, the lab must repeat EQA).
- Results from low coverage regions (<10X) can be excluded from final results.
- Results using another method (PCR or Sanger sequencing) can be provided with the final results when investigators believe the publicly available sequence data is not accurate.
- For exome or whole genome sequencing, the region chosen for reporting final results should have been sequenced previously at high depth (30X) in multiple populations and should be a region with low population minor allele frequencies.