

Genomic Data Analysis Network Centers Funding Application Questions and Answers

Q: Are the GDACs intended to be an analysis service for the CCG Program Office?

A: The governance of the GDAN will be similar to that of TCGA. There will be a Steering Committee overseeing the Program components as stated in the RFA. The GDAN is a cooperative agreement to do large-scale genomic analyses. The CCG Program Office will often request analyses from the GDACs and expect them to be done in a timely manner. An awardee cannot refuse a request.

Q: Do applicants have to address all CCG initiatives in their proposal or can they address capabilities that may apply to a wide range of different CCG projects?

A: The focus should be addressing the types of analyses that can be applied to CCG projects going forward, not a specific project.

Q: Do the pipelines developed by the GDACs have to be deployable at other institutions?

A: Applicants do not have to develop their own pipelines. They can use publically available tools and resources. For the Specialized GDACs, deployable pipelines are not a requirement, but if you do develop your own tools, they must be made publicly available.

Q: Can the Visualization GDACs provide analysis for CCG projects?

A: Yes, they can. The Visualization GDACs are intended to be more “outward looking” than “inward looking.” Visualization GDACs should have experience with integrating different data types and provide user-friendly tools for the research community, but those tools can be used for answering questions from the Analysis Working Groups (AWGs).

Q: What constitutes a first pass analysis that will provided by the Processing GDAC?

A: This refers to the standard types of analyses such as clustering, Significantly Mutated Gene lists, COSMIC runs, etc. The Processing GDAC should take every single sample and run it through their pipeline as a starting point for downstream analyses. Please refer to TCGA marker publications for examples of standard analyses.

Q: Are the Specialized GDACs expected to provide analyses for a certain type(s) of data or can they propose methodologies that apply to a wide range of different data types?

A: The types of data generated by CCG are “fixed” at the moment. For example, a proposed study of immunological profiles of different cancers would be out of the scope of the GDAN objectives because CCG does not plan to generate immunological data. Applicants must declare core competencies that are in line with the objectives of the GDAN RFA.

Q: Are commercial components of applications discouraged due to potential licensing fees?

A: Yes, they are. However, if there is a free version of a tool that is available commercially, it is recommended that the free version be used for the pipeline.

Q: Can other datasets such as ENCODE be used in analyses?

A: Yes, as long as the intent of using outside datasets is to inform CCG-specific projects and analyze the data generated by them. The primary goal of the RFA is to analyze CCG-produced data.

Q: Are the Visualization GDACs required to have a pre-existing tool that is already operational?

A: The tool has to be operational in some form. It does not need to be the final form, but these RFAs are not meant to be for data analysis, not tool development. Therefore, the tools MUST exist at the beginning of the award.

Q: How many Specialized GDACs will be funded?

A: The RFA authorizes up to 11 centers.

Q: Can clinical experts from the Federal Government be listed as collaborators on the application?

A: Federal Government employees can be listed as unfunded consultants/collaborators in the application. They cannot be listed as a PI in the Multi-PI component of the application, nor may any budgetary support be requested for them.

Q: Can the GCC protocol for generating Total RNAseq data be posted?

A: Yes, it will be posted in conjunction with this document and the webinar PowerPoint slides.

Q: Will there be both fresh frozen and FFPE cases?

A: Yes.

Q: For the “batch effects/integration” core competency, do they have be addressed together? Batch effects are not considered integration.

A: They can be addressed separately.

Q: Will there be extensive clinical annotations in new CCG projects?

A: Case reports will remain the same as they were for TCGA. The only difference is the extent of the follow-up data. For most of these projects, the cases will be collected from finished clinical trials. Thus the extent of follow up data will be on average 3+ years, depending on the cancer type.

Q: Will all cases have whole genome sequencing data?

A: The majority of cases will have low-pass whole genome sequencing data.

Q: Can animations be embedded within the application?

A: Web links cannot be part of the application, but embedded images can be, although it is the applicant’s responsibility to make sure they work after the submission is completed.

Q: Will the Processing GDAC deliver mutation calls?

A: The Genomic Data Commons (GDC) will process all of the BAM files and call mutations using an automated pipeline. The output will be VCFs.

Q: If a GDAC has new mutation caller, can that be integrated into the GDC?

A: A conversation can be started with the relevant GDC personnel about potentially integrating a new mutation caller.

Q: Can one of the specialized GDACs develop new mutation callers?

A: This is more in line with an investigative R01. Pure tool development is out of scope of the GDAN objectives.

Q: What are main differences between the Processing GDAC and the GDC?

A: The GDC will only do sequence alignment (BAM generation) and mutation calling. One of the core competencies of the Processing GDAC is to produce data on rearrangements, structural variants, etc. There is some overlap between the GDACs and the GDC.

Q: Is experimental validation of methods out of scope of the GDAN objectives?

A: Applicants need to show that the proposed tools will work at the beginning of the award period. As this RFA is NOT for tool development, validation of tools is out of scope.

Q: What about showing the use of new tools in preliminary results that have not been published?

A: Again, applicants should show evidence that the tools will work.

Q: Should a pipeline used in a Specialized GDAC be deployable outside of the developer's institution?

A: Yes, but this is not a requirement.

Q: Will the data generated by the Specialized GDACs be made available to the Visualization GDACs?

A: Yes, all results generated by the GDACs will eventually be made publically available at the GDC.

Q: Will germline calls be publically available or will they be under controlled-access?

A: The classification of controlled access data, the procedure for requesting controlled access data and types of controlled-access remain the same under NIH policy.

Q: Can the GDC pipelines be posted?

A: That is not available yet.

Q: The processing of nucleic acid data requires manual effort in checking BAM files. The cost of downloading and processing BAM files from the GDC can become quite expensive.

A: The level of funding is fixed as indicated in the RFA.

Q: What is the status of the NCI Cloud?

A: There is no single NCI Cloud at the moment, only Cloud Pilot projects that are temporary in nature. There are no current plans by NCI to create a free of charge compute space.

Q: Do indirect costs of third party providers count against the funding cap of the RFA?

A: The funding caps in the RFA are for *direct* costs. As such, indirect costs from the submitting institution as well as any third party participants do not count against the cap. The proposals must show budgets with total costs, clearly differentiating between direct and indirect costs.