RESOURCE SHARING POLICY (GENOMICS CORE)

All genomic and transcriptomic data generated with the help of the Core will be shared in accordance with the National Institutes of Health Genomic Data Sharing Policy (45 CFR 74.62). Non-human-subject sequencing data, including data from microbes, microbiomes, model organisms, and human cell lines purchased from commercial sources such as American Type Culture Collection (ATCC), as well as relevant associated data (e.g., phenotype and exposure conditions), will be deposited into a public repository within the required timeframe. The Core has used these repositories in the past, including Gene Expression Omnibus (GEO), Sequence Read Archive (SRA), and Mouse Genome Informatics (MGI). The GEO repository is the Core’s preferred database, and all genomics data generated with the help of the Core used in publications have been promptly deposited to the GEO database. All sequencing data will be made publicly available no later than the date of initial publication. If journal policy does not require it already; all users will be required to provide in the final version of each manuscript a web link or accession number for genomic and transcriptomic data generated with the help of the Core. At a minimum, all sequencing data will be deposited as raw files (in fastq format) including all associated data such as exposure, phenotype, etc. In addition, as required by the journal or database, processed data will also be deposited to the same repository alongside the raw data (for example, in bed, bedgraph, bw or bam formats). Any code used in bioinformatics data analyses will be made publicly available through shared depositories such as GitHub.

If the need arises, the Core (and any associated investigators) will apply to patent any intellectual property derived from in-house development of novel technologies or analyses methods, using the NIH Best Practices for the Licensing of Genomic Inventions and Section 8.2.3, Sharing Research Resources, of the NIH Grants Policy statement as guides. We will ensure that all genomic or genotype-phenotype data developed with NIH support will remain freely accessible in public databases.

Thus far, none of the projects processed by the Core, either from CoBRE-supported or external investigators, included human subject samples. However, should this need arise in the future, the Core will assist investigators to submit their large-scale human genomic data as well as relevant associated data (e.g., phenotype and exposure data) to an NIH-designated data repository in a timely manner, per NIH guidelines. For example, in the case of RNA expression profiling, this would be within six months after data submission is initiated or at the time of acceptance of initial publication, whichever occurs first. The Core will assist the investigators in submitting any information necessary to interpret the submitted genomic data, such as study protocols, data instruments, and survey tools.

If the Core receives human subject RNA or DNA samples for processing, the investigators submitting human subject samples will be required to strip the data of identifiers according to the Health Insurance Portability and Accountability Act (HIPAA) Privacy Rules. The de-identified data will be assigned random, unique codes by the

investigator, and the key to other study identifiers will be held by the submitting institution. Before starting a project with the Core, investigators submitting the samples will be required to confirm that informed consent for future research use and broad data sharing have been obtained, even if the human subject specimens are de-identified.

Additionally, the Core will assist the researchers in submitting the de-identified genomic data obtained from any human subject samples to NIH-designated data repositories according to the standards outlined in the HHS Regulations for the Protection of Human Subjects, to ensure that the identities of research subjects cannot be ascertained from the data. As an additional safeguard, we will assist investigators in requesting a Certificate of Confidentiality to prevent disclosure of any personally identifiable information they may hold.