Genomic Data Sharing

All genomic and transcriptomic data generated 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) and relevant associated data (e.g., phenotype and exposure conditions) will be deposited into a public repository within the required timeframe. We have extensive experience in using repositories in the past, including Gene Expression Omnibus (GEO), Sequence Read Archive (SRA), and Mouse Genome Informatics (MGI).

All sequencing data will be publicly available by the initial publication date. If the journal policy does not require it already, we will provide a web link or accession number for genomic and transcriptomic data generated in the final version of each manuscript. 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 be deposited to the same repository alongside the raw data (for example, in bed, biography, bw, or bam formats). Any code used in bioinformatics data analyses will be made publicly available through shared depositories such as GitHub.