



All data resulting from ADA-funded research that can be shared without compromising human subject protections must be shared to an approved open data repository within 6 months of publication or within 18 months of the conclusion of the funding period, if the study remains unpublished. Awardees are encouraged to use the repository most appropriate for the subject matter of the research conducted.

## Research Programs Approved Data-Sharing Repositories

[NIH Data Sharing Repositories](#) – a list of NIH-supported data repositories that make data accessible for reuse.

[EMBL Nucleotide Sequence Database](#) - Europe's primary nucleotide sequence resource. The main sources for DNA and RNA sequences are direct submissions from individual researchers, genome sequencing projects and patent applications. It is one part of the European Nucleotide Archive (ENA).

[COSMIC](#) - stores and displays somatic mutation information and related details and contains information relating to human cancers.

[Database of Genomic Variants](#) - aims to provide a comprehensive summary of structural variation in the human genome and provides a useful catalog of control data for studies aiming to correlate genomic variation with phenotypic data. The database is continuously updated with new data from peer reviewed research studies.

[The European Genome-phenome Archive](#) - designed to be a repository for all types of genotype experiments, including case control, population, and family studies. It includes SNP and CNV genotypes from array based methods and genotyping done with re-sequencing methods. This data may be either publicly available or limited access, depending on the design of the study.

[Protein Data Bank in Europe \(PDBe\)](#) - the EBI Protein Structure Database in Europe is a project for the collection, management and distribution of data about macromolecular structures, derived from the Protein Data Bank (PDB). It is one of the founding members of Worldwide Protein Data Bank (wwPDB).

[IntACT](#) - a freely available, open source database system and analysis tools for protein interaction data. All interactions are derived from literature curation or direct user submissions and are freely available.

[ArrayExpress](#) - a database of functional genomics experiments including gene expression where you can query and download data collected to MIAME and MINSEQE standards. Gene Expression Atlas contains a subset of curated and re-annotated Archive data which can be queried for individual gene expression under different biological conditions across experiments.

[PRIDE](#) - the PRIDE PRoteomics IDentifications database at EMBL-EBI is a centralised, standards compliant, public data repository for proteomics data. It has been developed to provide the proteomics community with a public repository for protein and peptide identifications together with the evidence supporting these identifications. PRIDE is also able to capture details of post-translational modifications.

[UK Data Archive \(UKDA\)](#) - a centre of expertise in data acquisition, preservation, dissemination and promotion and is curator of the largest collection of digital data in the social sciences and humanities in the UK.

[National Collection of Type Cultures \(NCTC\)](#) - a specialised laboratory located in the Central Public Health Laboratory, Colindale. It accesses, preserves and supplies authentic cultures of bacteria and mycoplasmas that are pathogenic to man or other animals that may occur in food or water and in hospital or health related environments and which can be preserved by freeze-drying.

[National Collection of Pathogenic Viruses](#) - a wide-ranging archive of well-characterised, authenticated human pathogens which will resource the supply of viruses, and materials derived from them, to the scientific community.

[Figshare](#) (general) - figshare allows users to upload any file format to be made viewable in the browser so that figures, datasets, media, papers, posters, presentations and file sets can be disseminated.