HDinHD: A Data Portal for Huntington's Disease Research
https://www.hdinhd.org/

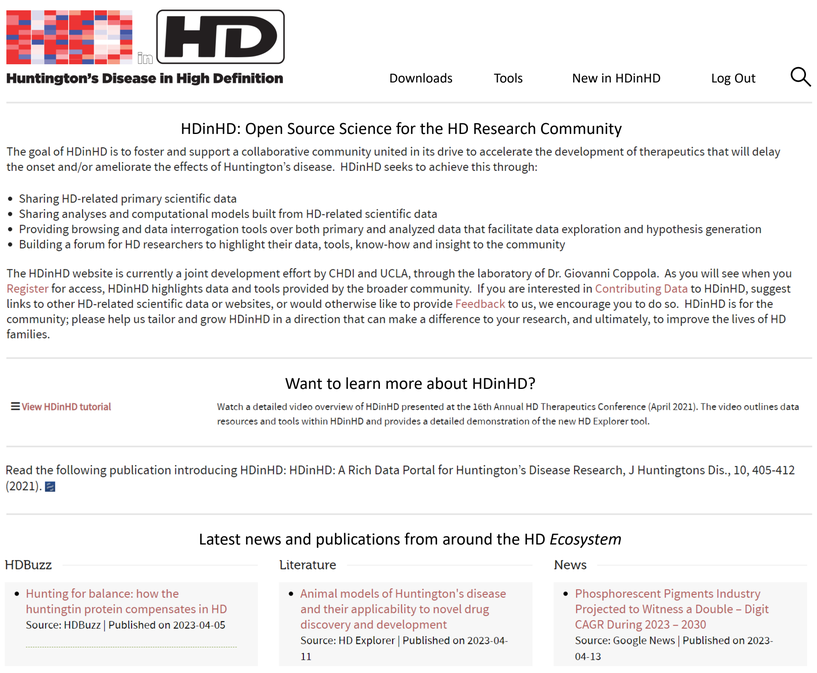
Viktoria Andreeva[3], Richard A. Blevins[2], Irina Murasheva[3], Kabenla Armah[3], Maksim Kastsiuchenka[3], Rob Gill[3], Jian Chen[1], Jim Rosinski[1], Liliana Menalled[1], David Howland[1], Dan Felsenfeld[1], Vahri Beaumont[1], Thomas F. Vogt[1], Jeff Aaronson[1]
[1]CHDI Management/CHDI Foundation, Princeton, NJ, USA. [2]Bridlewood Consulting LLC, Flower Mound, TX, USA. [3]Rancho Biosciences LLC, San Diego, CA, USA.

Abstract

HDinHD (Huntington's Disease in High Definition; HDinHD.org) is an open online portal for Huntington's Disease (HD) researchers. HDinHD presents an integrated view of HD experimental data and highlights a federated set of visualization and analysis tools developed independently by HD scientists across the community. Researchers can interactively explore a richly-connected set of experimental data, visualize analytical results or download datasets locally to incorporate bulk data into their own databases and computational pipelines.

The HDinHD team continues to monitor the literature and community omics repositories to identify emerging HD experimental data. New HD studies are curated and analyzed according to established methods and pipelines and integrated into the HDinHD environment. HDinHD endeavors to be responsive to community needs. Input is welcomed and may be submitted either through the Feedback link on the HDinHD website or by contacting CHDI directly.

HDinHD Home Page



HDinHD home page content includes a video tutorial, RSS feeds from HDBuzz, HD-related literature aggregated from PubMed, bioRxiv and medRxiv and an HD-related news feed. Downloads, Tools, and New in HDinHD sections are available as independent tabs from the home page to registered users. Prospective users may register for an account directly from the homepage to gain full access.

Downloads

Striatum Disease Signature	Manuscript describing generation of molecular disease signatures in HD mice and supplemental files detailing results [2].
Mouse Allelic Series	Raw, processed and analyzed molecular and behavioural data from the Mouse Allelic Series project.
GWAS Studies	Topic reports for genes implicated by early GeM-HD results.
DNA Repair & Handling	Topic report plus visual and computable DNA repair pathways.
Causal Modeling Results	Simulation and other results from a series of causal models built from Mouse Allelic Series molecular and behavioural data.
Curated HD Datasets	Independent slices of HD experimental data, including **HDSigDB**, underlying HD Explorer Tool.

Downloads: HDSigDB

To provide rich functional context for HD gene set enrichment analysis, we developed an HD-relevant gene set library called **HDSigDB**. The core of **HDSigDB** was derived from curation and analysis of HD and triplet-repeat expansion disease studies deposited in GEO, ArrayExpress, and PRIDE. Additional sources of gene sets include selected PubMed articles (publication-based) and DNA Damage Response pathways (CHDI reports). The lists of genes available in human and mouse symbols and Entrez gene IDs.

HDSigDB is available within HDinHD Downloads and integrated within HD Explorer. **HDSigDB** is also included in the Enricher, a gene set enrichment analysis package from Ma'ayan Lab (Mt. Sinai) (https://maayanlab.cloud/Enrichr/) [3-5].

HDSigDB Gene Sets

Gene Set Category	Gene Set Sub-category	# of Gene Sets
Disease and Disease Model Signatures	Huntington's disease	1,797
	Ataxia (including Friedrich's, SCA1, SCA2, SCA3, SCA6, SCA7, SCA17)	487
	Muscular Dystrophy (including Becker, Duchenne, tibial, myotonic, facioscapulohumeral)	191
	Other diseases	131
Non-Disease Signatures	Tissue or cell type signatures (neurons, astrocytes, etc.)	778
	Pathways (DNA repair pathways, perturbation effects in WT, etc.)	256
	Protein localization (cytoplasm, nucleus, etc.)	15
Other		17

Federated Set of HD Tools Authored by the Community

	HD Explorer	Integrated network of HD experimental data curated and analyzed from the literature, community omics repositories and newly-released internal CHDI reports.
GeM-HD Consortium	**GeM MOA SNP Viewer**	Summary findings from Huntington's disease genome-wide association studies that seek out genes influencing the pathogenesis and expression of Huntington's disease.
GeM-HD Consortium	**GeM Euro 9K**	Visualization tools and summary results of a genome-wide association study to identify genetic modifiers of Huntington's disease.
	ASViewer	Visualization of Q-length and age dependent gene and protein expression data from brain and peripheral tissues of the Mouse Allelic Series.
Khakh Lab (UCLA)	**Adult Astrocyte RNAseq Explorer**	Visualization tool providing Astrocyte gene expression profiles across brain regions and HD disease models.
Schaab (Evotec)	**HD Proteome Base**	Proteomics query tool displaying differential expression data from brain and peripheral tissues of the Mouse Allelic Series as well as baseline proteomic and phosphoproteomic data from the R6/2 mouse model.
Neri Lab (INSERM)	**BioGemix Suite**	Browsable knowledgebase of integrated HD animal model data using precision machine-learning and 3D-visualisation of RNA-seq data in brain structures of HD model mice.
Ma'ayan Lab (Mt. Sinai)	**Enricher**	Gene set enrichment analysis tool operating over a large, diverse collection of gene set libraries including HDSigDB, a gene set library containing HD and HD-related gene sets.
Yang Lab (UCLA)	**CoExMap Viewer**	Visualization tool allowing exploration of results and underlying data of a large-scale Weighted Gene Co-expression Network Analysis (WGCNA) of hundreds of samples from intact mouse striatum at 6-month of age as well as from gene set enrichment analysis of transcriptomic signatures of differentially expressed genes from 52 heterozygous HD knockout mice and wildtype controls.

Tools: CoExMap Viewer

CoExMap Viewer is now available as a federated application within HDinHD's Tools section [6]. CoExMap Viewer offers two major interactive components:

1. A Table Browser that allows users to sort striatal gene signatures, cell type, and/or diurnal markers based on their enrichment in any CoExMap module as well as the converse, to sort the CoExMap modules by relevance to any particular gene signature, cell type or diurnal marker.
2. A UMAP (uniform manifold approximation and projection) visualization that depicts a 2-dimensional projection of co-expression modules that reflects intramodular connectivity. Users can visualize the location of their own set of genes on the UMAP, and gain insight based on functional annotation of distinct regions within the UMAP.

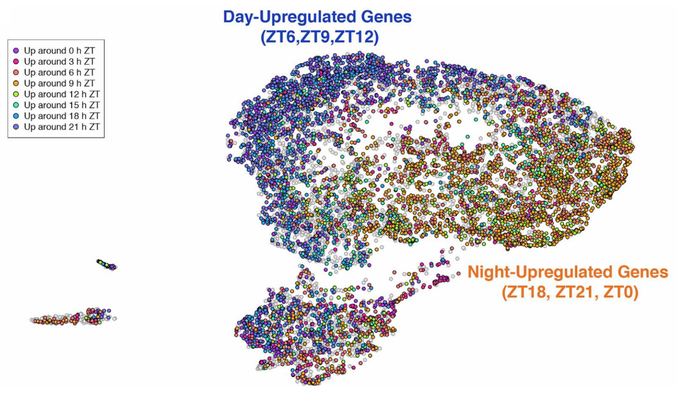
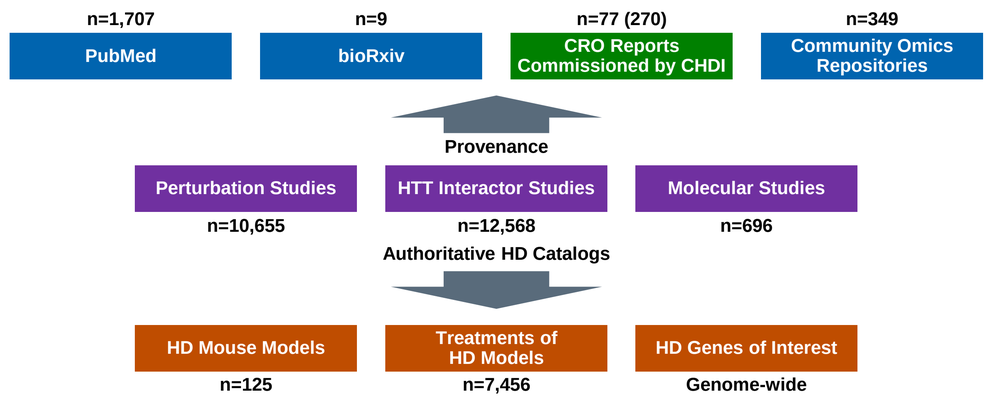


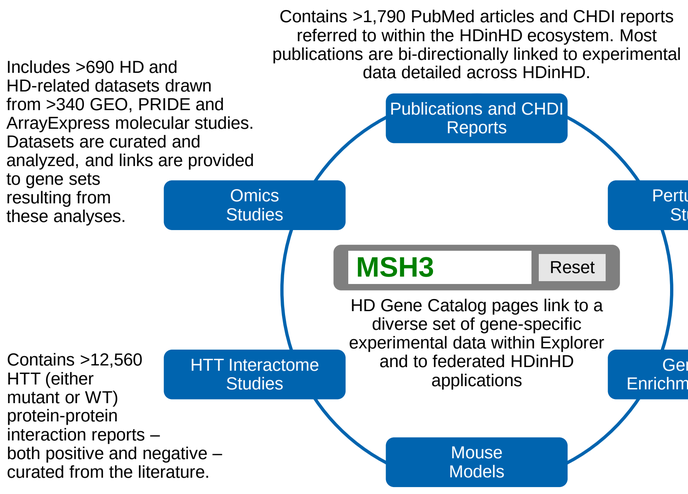
Figure 3D from [6]

Tools: HD Explorer

HD and HD-related Experimental Data Curated & Analyzed from Internal and External Sources. Shared HD Catalogs allow facile pivoting on Mouse Models, Treatments and Genes/Targets.



HD Explorer Entry Portals



Each box on the Explorer entry circle as well as the central gene name search box, provide distinct portals into the integrated HD Explorer application. Once inside a portal, users can pivot via a rich set of semantic links to explore related HDinHD federated tools or adjacent HD Explorer data sections. For example, after entering a gene name, users check on the expression of that gene/protein in the Mouse Allelic Series and can visit sets of experimental studies performed on that gene within the HD context.

Component datasets are available on the HDinHD Downloads tab for labs who wish to incorporate HD Explorer batch data to facilitate internal data mining efforts.

Acknowledgements

HDinHD is funded and developed by CHDI Foundation, Inc., a nonprofit biomedical research organization exclusively dedicated to collaboratively developing therapeutics that will substantially improve the lives of those affected by Huntington's disease. HDinHD was launched in 2015 in partnership with the laboratory of Giovanni Coppola (UCLA). Colleagues at Rancho BioSciences contributed data curation, data analysis, data modeling and software/data engineering support, and Bridlewood Consulting contributed solutions architecture, systems and software engineering support. CHDI thanks the investigators who have kindly contributed to HDinHD's federated set of community-developed tools.

References

1.Aaronson J et al, *J Huntington's Dis* 2021,10(3):405-412.
2.Obenauer J et al, bioRxiv, https://doi.org/10.1101/2022.02.04.479180
3.Chen EY, et al, *BMC Bioinformatics* 2013, 128(14).
4.Kuleshov MV et al, *Nucleic Acids Research* 2016, gkw377.
5.Xie Z et al, *Current Protocols* 2021, doi:10.1002/cpz1.90.
6.Wang N et al, *Neuron* 2022, 110(20):3318-3338