

dbGaP Study Release Notes



Release Notes for NIMH SCAP-T, phs000833.v1.p1 “Single Cell Analysis Program - Transcriptome (SCAP-T)”

For any questions or comments, please contact: dbgap-help@ncbi.nlm.nih.gov.

February 17, 2015 Version 1 Data set release date

2015-02-17

Version 1 Data set release for NIMH SCAP-T now available

In the parent study and substudy model, the top-level study phs000833.v1.p1 NIMH SCAP-T Study contains the Subject Consent file, Subject Sample Mapping file, and subject phenotypes shared among all centers (UCSD, UPenn, and USC).

Substudies have been created to house any center-specific phenotype variables and molecular genetic data, specifically RNASeq data and expression levels derived from RNA sequences, on SCAP-T participants. The substudies for this current release include:

- phs000834.v1.p1 Single Cell Analysis Program - Transcriptome (SCAP-T) (UCSD)
http://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs000834
- phs000835.v1.p1 Single Cell Analysis Program - Transcriptome (SCAP-T) (U. Penn)
http://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs000835
- phs000836.v1.p1 Single Cell Analysis Program - Transcriptome (SCAP-T) (USC)
http://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs000836

Approved authorized users for phs000833.v1.p1 will be able to download data from the top-level study and all substudies.

**There are no overlapping subjects in the 2 consent groups listed below.

Consent group 1 (c1): General Research Use (GRU)

	phenotype	SRA/expression
total subjects	6	6
total samples	512	512

phs	substudy	subjects	samples
phs000834.v1.p1	SCAP-T UCSD	1	497
phs000835.v1.p1	SCAP-T UPenn	0	0
phs000836.v1.p1	SCAP-T USC	5	15

Consent group 2 (c2): General Research Use (NPU) (GRU-NPU)

	phenotype	SRA/expression
total subjects	8	8
total samples	185	185

phs	substudy	subjects	samples
-----	----------	----------	---------

dbGaP Study Release Notes



phs000834.v1.p1	SCAP-T UCSD	0	0
phs000835.v1.p1	SCAP-T UPenn	8	185
phs000836.v1.p1	SCAP-T USC	0	0

Additional notes for molecular data

1. Sample-level data are provided by three sub-studies, phs000834.v1.p1, phs000835.v1.p1 and phs000836.v1.p1. They are organized and accessioned as phg000535.v1, phg000536.v1 and phg000537.v1 respectively.
2. Please see the components marked as “sample-info” for sample-subject mapping, sample-consent status and sample-data file mapping of each sub-study.
3. Expression data, which were derived from RNA sequencing, are split on samples by submitters. Two sequence-count files, labeled as “intronic” and “exonic”, are associated to one sample. They are packed into “expression-data-indfmt” components for each accessioned folder.

Authorized Access (Individual Level Data and SRA Data)

Individual level data and SRA sequencing data are available for download through the dbGaP Authorized Access System upon approval of the Data Access Request (DAR):

- <http://dbgap.ncbi.nlm.nih.gov/aa/wga.cgi?login=&page=login>

Public FTP site (Summary Level Data Only)

All data tables, data dictionaries, and documents will be housed under one directory for ease of downloading. The data_dict filenames have an added study version number (phs#.v#) and deleted participant set number (p#) from the table accession (pht#.v#). The var_report filenames have an added study version number (phs#.v#). In the var_report files, variables contain version numbers (phv#.v#) and summaries were created for each consent group (c#). These FTP files are available at:

- <ftp://ftp.ncbi.nlm.nih.gov/dbgap/studies/phs000833/phs000833.v1.p1>

*****Penn and USC images are available for download under the public FTP “documents” directory.**