VIRAHEP-C Ancillary Studies Data Archive

The VIRAHEP-C Ancillary Studies Data Archive contains a study protocol, the raw data collected during the study, and analysis files.

This new version of the data archive includes the datasets for an additional Genotyping Ancillary Study, which is not recorded in the original study protocol.

VIRAHEP-C Ancillary Studies

The VIRAHEP-C study includes the main study and the following five ancillary studies:

- Viral Genetics
- Patient Genetics
- Interferon Signaling
- Immunology
- Genotyping

This document describes the complete archive for the ancillary studies. Ancillary Study "Version 1.0" contains data from the viral genetics, patient genetics, and interferon signaling studies. This new Ancillary Study "Version 2.0" contains new additions of Immunology and Genotyping studies. The archive for the main study is also available from the NIDDK Central Repository.

The ancillary studies archive is organized into folders corresponding to each of the studies. It also contains a document, "Virahep-C_Protocol_Ancillary.pdf", describing the protocol for the ancillary studies. The information in this document was extracted from the protocol for the main study, which is included in the main study archive.

1. Viral Genetics Ancillary Study

Principal Investigator: John Tavis, PhD
St. Louis University, St. Louis

The archive for the Viral Genetics study contains summary data and the raw data. The raw data is 4 GB.

- Summary data includes the following spreadsheets:
    - "Pre-treatment_sequencing.xls"
    - "Post-treatment_sequencing.xls"
- The raw data is available upon request due to its large data size. There is a "Readme" file describing the location of 4GB raw chromatogram files.

2. Patient Genetics Ancillary Study

Principal Investigator: Huiying Yang, MD, PhD
Cedars-Sinai Medical Center, Los Angeles

The archive contains 5 datasets organized into 5 folders. Table 1 lists the names of the dataset(s) in each

folder and provides a brief description. This information is also provided in the dataset description document, "Dataset_description_genetic.pdf". This study is published in Hepatology *2007*[1] .

Table 1: Dataset Description for the Patient Genetics Ancillary Study

Folder	Dataset	Description
01_Allele_data_197_genes	Proc_contents_vhepsnps03292007.doc	Allele data on 1082 SNPs (2 copies each) on 197 genes.
	vhepsnps03292007.sas7bdat	
02_Population_Marker	poplmark.sas7bdat	Population, or "ancestry-informative" markers for genetic Virahep-C patients. For more information on their use, see reference: Yee et al. Mxyovirus-1 (Mx1) and protein kinase (PKR) haplotypes and fibrosis in chronic HCV, 2007.
03_Haplotype	Haplotype datasets:	Haplotype refers to "the combination of SNPs within a gene that tend to be inherited together as a block. Haplotype data are received at the chromosome level. All haplotypes were determined using the PHASE® software. For details on haplotype construction, see reference: Yee et al Mxyovirus-1 (Mx1) and protein kinase (PKR) haplotypes and fibrosis in chronic HCV, 2007.
	ifi35	
	ifnar1	
	ifnar2	
	ifng	
	il10	
	ip_10_cxcl10_	
	irf1	
	irf7	
	irf9_isgf3g_	
	isg15_g1p2_	
	mx1	
	mx2	
	mxa	
	oas2	
	oas3	
	oas_1	
	oasl	
	pkr	
	stat1	
	stat2	
	tgfb1	
	hla_haplotype	
	proc_contents_haplotype.pdf	
04_HLA_Allele	Proc_contents_hla_allele.pdf	Allele data on HLA class I and class II genes, i.e. HLA genes A, B, C, and DPA1, DPB1, DRB1, DQB1, DQA1.
	hla_allele.sas7bdat	
05_Gene_map	GENE_map.xls	GENE_map.xls Provides SNP, chromosome number, gene name, and location for the genes included in the genetic study.

[1] Yee L, Tang Y, Kleiner D, Wang D, Im K, Wahed A, Tong X, Rhodes S, Su X, Whelan R, Fontana R, Ghany M, Borg B, Liang T, Yang H. Myxovirus-1 and protein kinase haplotypes and fibrosis in chronic hepatitis C virus. *Hepatology* Jul 2007 46(1):74-83

3. Interferon Signaling Ancillary Study
Principal Investigator: Milton Taylor, PhD
Indiana University, Bloomington

The archive for the Interferon Signaling study contains summary data listed below. The study is published in Journal of Virology 2007[2].

- The "microarray_data.sas7bdat" file contains the expression level data created using the Affymetrix microarray chip HG-U133A. The data can be linked to the clinical data in the Main Study by VHCID and TMPT. Annotation for "Gene name" can be found in the annotation file http://www.affymetrix.com/support/technical/byproduct.affx?product=hgu133.
- "microarray_format_file.sas" is a format file for the SAS data file above.
- "proc_contents_microarray_dataset.doc" is a content description file for the SAS data file above.

4. Immunology Ancillary Study
Principal Investigator: Hugo Rosen, PhD
University of Colorado (previously Portland VAMC)

The archive for the Immunology study contains data files listed below.

4.1 CD4 ELISPOT – This study was published in Hepatology. 2007 Aug;46(2):350-8. Whole PBMCs from patients in Virahep-C were stimulated with recombinant HCV proteins as described in methods and IFN-? ELISPOT production was assessed.
- elispot.sas7bdat
- sumarea.sas7bdat
- sumarea_proc_content.pdf

4.2 PD-1 – Flow cytometry was used to measure programmed death-1 (PD-1) expression on T cells, Natural Killer (NK) and NKT cells. For HCV-specific CD8+ T cells, we used tetramers to enumerate and phenotype. Comparisons were made between races, according to response to antiviral therapy, and pre and post-treatment. Results were published in J Immunol. 2008 Mar 15;180(6):3637-41.
- Dataset_description_pd1.pdf
- pd1pent.sas7bdat
- pd1time.sas7bdat

4.3 Treg - Dataset contains the percentage of Regulatory T cell (Treg) expressing CD152, CD62L (positive/negative), FOXp3, CD69, LAG-3 and CD45Ro in addition to overall Treg frequencies.
- Treg_data_description.pdf
- Treg_format.sas
- tregdata.sas7bdat

4.4 Dendritic Cell study - Expression level measurements on Plasmacytoid (pDC) and myeloid Dendritic Cells (mDC) for 66 participants from the Virahep-C cohort on a group of genes. The

[2] Taylor MW, Tsukahara T, Brodsky L, Schaley J, Sanda C, Stephens MJ, McClintick JN, Edenberg HJ, Li L, Tavis JE, Howell C, Belle SH. Changes in gene expression during pegylated interferon and ribavirin therapy of chronic hepatitis
C virus distinguish responders from nonresponders to antiviral therapy. J Virol. 2007 Apr;81(7):3391-401.

study was published in Gut 2009[3].

- Dendritic_Cell_data_description.pdf
- dcell.sas7bdat
- VHC DC Data.xlsx

4.5 NK Cell study – Using flow cytometry, the phenotype of natural killer cells was assessed in over 100 patients prior to treatment, including 14 activating or inhibiting receptors and correlations were made with viral kinetics and sustained virologic response. The manuscript was accepted for publication in Hepatology in June 2011.

- VHC_NK_data_dictionary.doc
- VHC_NK.xls

5. Genotyping Ancillary Study

Principal Investigator: Janardan Pandey, PhD
Medical University of South Carolina, Charleston

The Genotyping data is generated from the "Immunoglobulin Allotypes and Response Rates to Antiviral Therapy for Hepatitis C" ancillary study. This study used PCR-RFLP and direct DNA sequencing methods for genotyping 319 Virahep-C subjects available from the NIDDK Repository. For QC measures, repeat typing of the markers was performed on 5% of the samples. Unblinding of the quality control samples was performed only after both samples had been completely analyzed. Some coded samples were characterized by two different techniques; both provided identical results (Hum Immunol, 2002[4]).

The dataset contains summary data listed below.

- The "genotypes.txt" file contains the genotyping result data for 10 markers.
- The "readme.txt" file provides information about the genotypes data file.

[3] Mengshol, J. A., Golden-Mason, L., Castelblanco, N., Im, K., Dillon, S., Wilson, C. C., Rosen, H. R. for the Virahep-C Study Group, Impaired Plasmacytoid Dendritic Cell Maturation and Differential Chemotaxis in Chronic HCV: Associations with Antiviral Treatment Outcomes, 2009 Gut 58(7): 964-73.

[4] Pandey JP, Frederick M; ACCESS Research Group. A Case Control Etiologic Study of Sarcoidosis. TNF-alpha, IL1-beta, and immunoglobulin (GM and KM) gene polymorphisms in sarcoidosis. Hum Immunol. 2002 Jun;63(6):485-91.