

SickleGen and Sickle Cell Genomics

Martin H. Steinberg for the SickleGen Consortium

Department of Medicine, Boston University School of
Medicine, Boston, MA



SickleGen PIs

Paola Sebastiani Clinton Baldwin	Boston University
Michael DeBaun	Washington Univ., Vanderbilt
James Casella	Johns Hopkins
Victor Gordeuk	Howard
Mark Gladwin	Univ. of Pittsburgh
Marilyn Telen	Duke
Greg Kato James Taylor	NHLBI
Carolyn Hoppe	Children's Hospital Oakland
Abdulrahman Alsultan	King Saud Univ., Riyadh
Amein Al-Ali	King Faisal Univ., Dammam

Sharing: Summary data on dbGaP, individual data shared among SickleGen and outside investigators, individual data on dbGap awaiting approvals

Cohorts Already Genotyped with Phenotypes and Biological Samples

Center/PI/Grant	Source/ Cases
BU/Steinberg/HL068970; HL78681	CSSCD/1500
BU/Steinberg/HL70735	MSH/280
BU/Klings/K23HL79003	BMC/200
Children's Hospital Oakland/Hoppe/U54 HL70583	CSCC C-data Gen-Phen project/600
Pittsburg/Gladwin/ HHSN 2680061784C	Walk-PHaSST/635
Duke/Telen/HL079915; HL68959	Outcome and Modifying Genes in Sickle Cell Disease/PH in SCD/500
Howard/Gordeuk/HL79921	PUSH/510
Vanderbilt/Hopkins/DeBaun, Casella/NINDS U01 NS042804	SITT trial/1130 collected (additional collections ongoing)
NIH/NHLBI/Taylor	NIH PH Study/530
King Faisal/Al-Ali/HL068970	HbF in Saudi Arabia/200

SickleGen Database

We created a database with free access to investigators who contributed genetic and phenotypic data

The SickleGen Database: A database for genotype-phenotype studies in sickle cell anemia

home · about · help

QUESTIONNAIRES | SNP DATA | WORKSPACE | DOCUMENTATION | SEARCH

HOME

Welcome to the SickleGen database.

Use menu on top to navigate.

SickleGen database is a collection of phenotype and genotype data from different databases brought together to study sickle cell anemia. Current list of databases include CSSCD, MSH, C-Data, BU-PH. Main features of the database are outlined below. For more complete help manual please follow the help link on top of this page.

Questionnaires tab contains data from various data sets produced from patient questionnaires.

Each data set can be sorted by any field by clicking its name. Each set can be filtered and saved as a group of ids (see filtering for more information).

Set name and database name are printed on top of the page, followed by filter box, followed by page information and navigation are finally by top 100 records returned by the database.

Questionnaires Filtering

QUESTIONNAIRES | SNP DATA | WORKSPACE | DOCUMENTATION | SEARCH

QUESTIONNAIRES This section of the database provides ways to search and filter phenotypes information.

- Select phenotype table from the menu on the right.
- Enter the filtering criteria to narrow down your search.

>> FORM: "CSSCD.R01"

Filtering Options:

(1149) DATE OF VISIT (FOIDATE) greater than 1979-10-17 APPLY

(7) FORMID = DLR

Before Filters: 3,770; After Filters: 1,090

showing 1-100 out of 1,090

1 2 3 4 5 6 >>

goto page 1 out of 11 90

DOWNLOAD SAVE AS

ANONYMIZED ID #	PAGE NUMBER OF FORM WITHIN HOUSEHOLD	ENTRY CLERK INITIALS	DATE FORM ENTERED	DATE OF LAST EDIT	DATE OF LAST EDIT	EDIT STATUS	VERSION DATE	DATE OF VISIT	FORM COMPLETED BY	FORM NUMBER RESPONSE
107591	1	DLR	1980-07-18	1980-10-15	1980-10-15	6	1979-06-11	1979-10-17	3	2
181364	1	DLR	1980-07-25	1980-10-15	1980-10-15	6	1979-06-11	1979-10-18	4	3

SickleGen Database

Five studies included. The database includes all data from the 3 phases of the CSSCD study and derived variables



The screenshot shows the SickleGen Database website. At the top, there is a header with the title "The SickleGen Database:" and a subtitle "A database for genotype-phenotype studies in sickle cell anemia". Below the header is a navigation bar with links for "home", "about", and "help". A dark blue navigation bar contains the main menu items: "QUESTIONNAIRES", "SNP DATA", "WORKSPACE", "DOCUMENTATION", and "SEARCH". On the left side, there is a sidebar with the heading "QUESTIONNAIRES" and a list of study names: "CSSCD", "C-DATA", "MSH_Clinical", "MSH_Ext_1", "MSH_PFU", "BU-PH", and "Howard Group". The main content area on the right contains a paragraph explaining that this section provides ways to search and filter phenotypes information, followed by two bullet points: "Select phenotype table from the menu on the right." and "Enter the filtering criteria to narrow down your search."

The SickleGen Database:
A database for genotype-phenotype studies in sickle cell anemia

[home](#) • [about](#) • [help](#)

[QUESTIONNAIRES](#) | [SNP DATA](#) | [WORKSPACE](#) | [DOCUMENTATION](#) | [SEARCH](#)

QUESTIONNAIRES

- ▶ CSSCD
- ▶ C-DATA
- ▶ MSH_Clinical
- ▶ MSH_Ext_1
- ▶ MSH_PFU
- ▶ BU-PH
- ▶ Howard Group

This section of the database provides ways to search and filter phenotypes information.

- Select phenotype table from the menu on the right.
- Enter the filtering criteria to narrow down your search.

SickleGen Database

Site-Adjusted Variables and Hemolytic Index in CSSCD

QUESTIONNAIRES | SNP DATA | WORKSPACE | DOCUMENTATION | SEARCH

QUESTIONNAIRES

This section of the database provides ways to search and filter phenotypes information.

- Select phenotype table from the menu on the right.
- Enter the filtering criteria to narrow down your search.

>> FORM: "CSSCD.C_HEMO_INDEX"

(help) **Filtering Options**

AGE_BIL

Before Filters: 3,250; After Filters: 3,250

showing 1-100 out of 3,250

1 2 3 4 5 6 >>

goto page out of 33

ANONID	TABLE_INDEX	SEX	AGE_BIL	LOG_BIL	AGE_LDH	LOG_LDH	AGE_RETIC	LOG_RETIC	AGE_SGOT	LOG_AGE	PC1	PC2	PC3	PC4
181384	1	1	33.60985626	2.29797448	35.56468173	6.684111803	38.07865982	3.230804396	38.0698152	3.954507817	-4.148772948	0.47343058	0.773862435	0.889602171
181449	2	2	0.963723477	0.70902053	0.963723477	6.135781318	1.393568051	2.611172126	3.119780972	4.38321343	-1.707215394	-1.652272233	-0.312777401	-0.915394079
198869	3	1	51.72621492	2.151180639	52.76112252	5.9821715	52.72279261	2.966818263	52.76112252	3.165264055	-1.237116818	2.619940318	0.401280312	0.273731743
215820	4	1	17.42642026	1.101275406	18.45448323	6.08221891	18.34086243	2.708383479	18.45448323	3.769537357	-1.170889919	0.08632969	0.334885331	-0.434726933

SickleGen Database

Phenotype Data

The screenshot shows a web browser window displaying the SickleGen Database interface. The browser's address bar shows the URL: `http://lobstah.bu.edu/gdd/ph.php?db=Howar`. The page title is "The SickleGen Database: A database for genotype-phenotype studies in sickle cell anemia".

The navigation menu includes: [home](#) - [about](#) - [help](#). The main menu has: [QUESTIONNAIRES](#) | [SNP DATA](#) | [WORKSPACE](#) | [DOCUMENTATION](#) | [SEARCH](#).

The "QUESTIONNAIRES" section is active, displaying the text: "This section of the database provides ways to search and filter phenotypes information." It includes a list of questionnaire groups: Howard_Group, HOWARD_PUSH, and PITTSBURG. The selected form is "Howard_Group.PITTSBURG".

The "Filtering Options" section shows a dropdown menu with "ac_loc1" selected and an "APPLY" button.

Summary statistics: "Before Filters: 570; After Filters: 570" and "showing 1-100 out of 570".

Page navigation: "1 2 3 4 5 6 >>" and "goto page 1 out of 6 go".

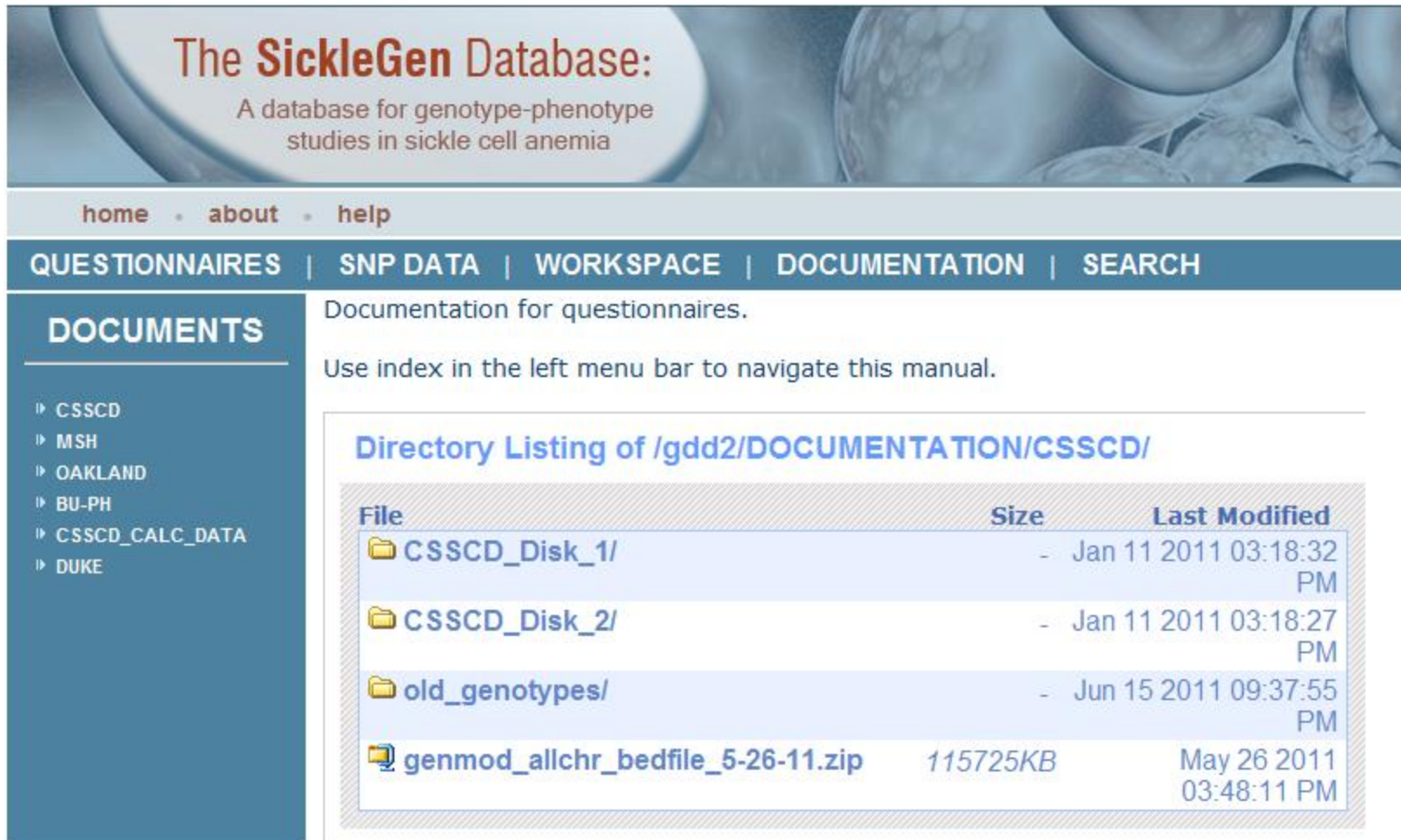
Buttons: "DOWNLOAD" and "SAVE AS".

pid	site_id	str	trtl	gender	age	scgeno	datastr	phase	seqno	site	project	edc_key	parent_key	asmt dt	asmt da	asmt mo	asmt yr	dob dt	dob da	dob mo	dob yr
10010	01			Male	52.9000015258789	SB+ (thalassemia)	DEMO	1000	000	Albert Einstein College of Medicine	Screening	121218	121217	2008-04-28	28	4	2008	1955-05-28	28	5	

Zoom level: 90%

SickleGen Database

Genotype Data



The SickleGen Database:
A database for genotype-phenotype studies in sickle cell anemia

home • about • help

QUESTIONNAIRES | SNP DATA | WORKSPACE | DOCUMENTATION | SEARCH


DOCUMENTS

- ▶ CSSCD
- ▶ MSH
- ▶ OAKLAND
- ▶ BU-PH
- ▶ CSSCD_CALC_DATA
- ▶ DUKE

Documentation for questionnaires.

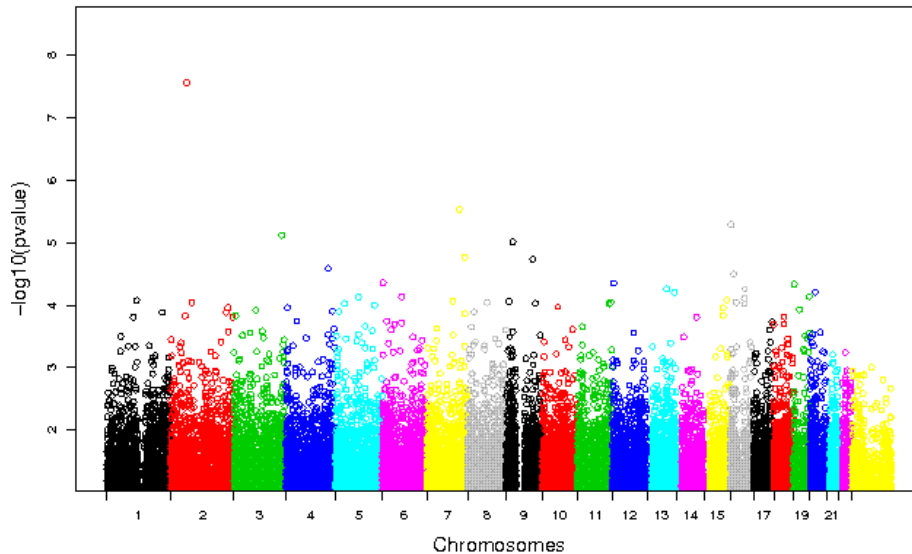
Use index in the left menu bar to navigate this manual.

Directory Listing of /gdd2/DOCUMENTATION/CSSCD/

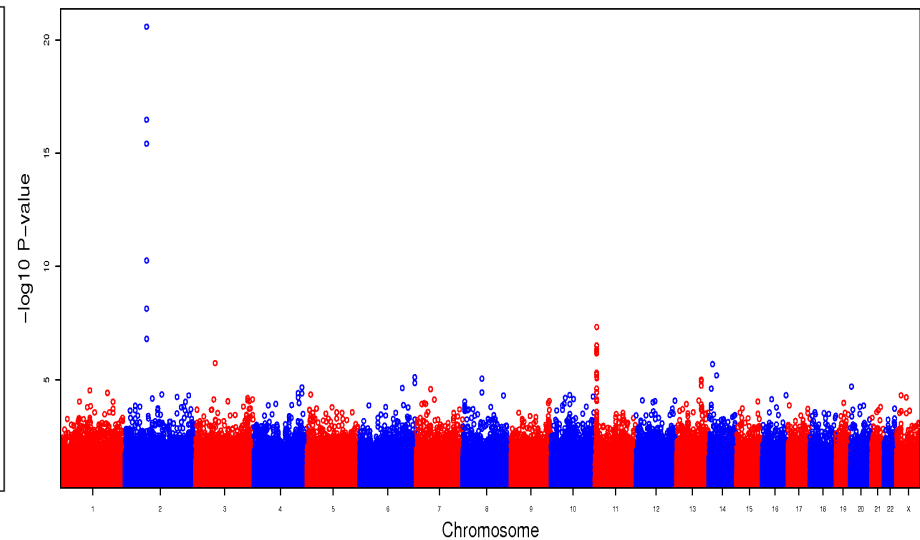
File	Size	Last Modified
 CSSCD_Disk_1/	-	Jan 11 2011 03:18:32 PM
 CSSCD_Disk_2/	-	Jan 11 2011 03:18:27 PM
 old_genotypes/	-	Jun 15 2011 09:37:55 PM
 genmod_allchr_bedfile_5-26-11.zip	115725KB	May 26 2011 03:48:11 PM

SickleGen GWAS Studies (HbF)

MSH HbF Sex Adjusted



CSSCD HbF Sex Adjusted



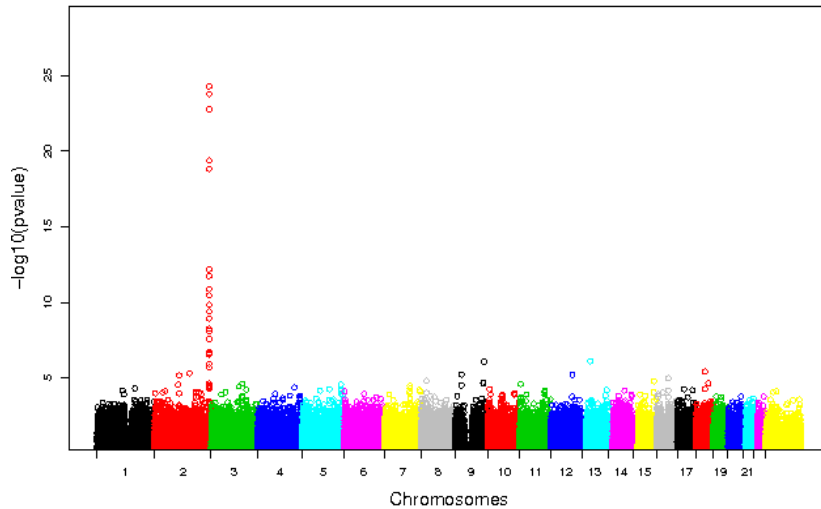
Sedgewick, A.E., Timofeev, N., Sebastiani, P., So, J.C.C., Ma, E.S.K., Chan, L.C., Fucharoen, G., Fucharoen, S., Barbosa, C.G., Vardarajan, B., Farrer, L.A., Baldwin, C.T., Steinberg, M.H., Chui, D.H.K. *BCL11A* is a major HbF quantitative trait locus in three different populations with β -hemoglobinopathies. *Blood Cells Mol. and Dis.* 41: 255, 2008.

Solovieff, N., Milton, J.N., Hartley, S.W., Sherva, R., Sebastiani, P., Dworkis, D.A., Klings, E.S., Farrer, L.A., Garrett, M.E., Ashley-Koch, A., Telen, M.J., Fucharoen, S., Ha, S.Y., Li, C.K., Chui, D.H.K., Baldwin, C.T., Steinberg, M.H. Fetal hemoglobin in sickle cell anemia: Genome-wide association studies suggest a regulatory region in the 5' olfactory receptor gene cluster. *Blood* 115: 1815, 2010.

SickleGen GWAS Studies (others)

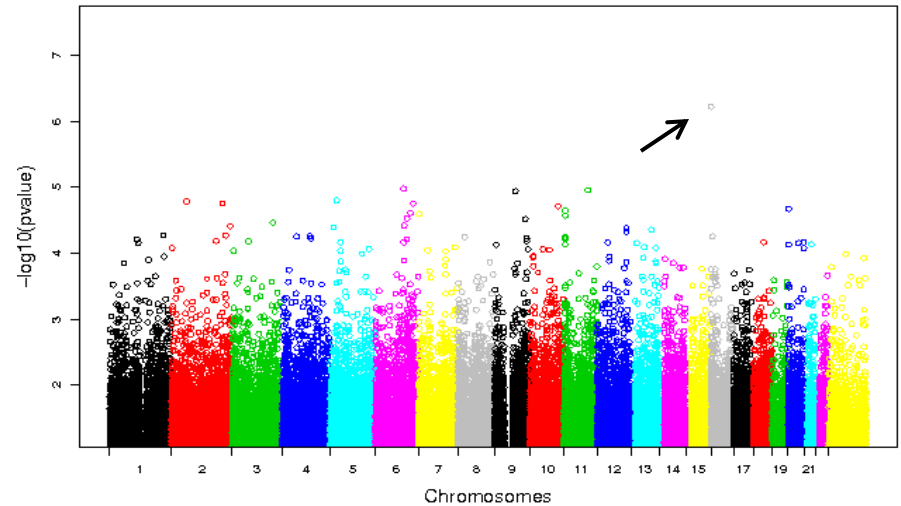
Bilirubin: *UGT1A1*

$-\log_{10}$ pvalue – Bil Additive age,sex Adjusted MAF>0.01



Hemolysis: *NPRL3*

$-\log_{10}$ pvalue – pc1 Additive age,sex Adjusted MAF>0.05



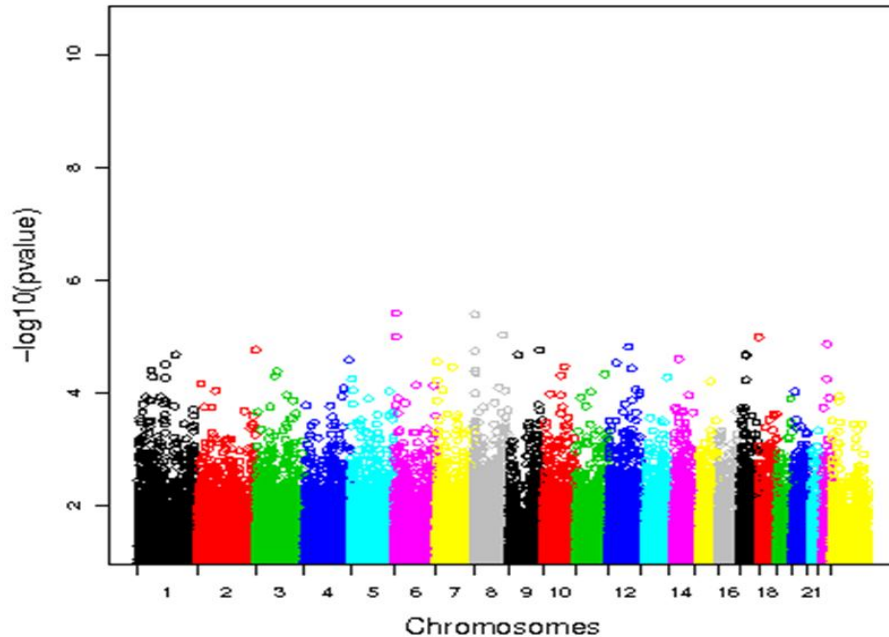
Milton, J.N., Sebastiani, P., Solovieff, N., Hartley, S.W., Bhatnagar, P., Arking, D.E., Dworkis, D.A., Casella, J.F., Barron-Casella, E., Bean, C.J., Hooper, W.C., DeBaun, M.R., Garrett, M.E., Soldano, K., Telen, M.J., Ashley-Koch, A., Gladwin, M.T., Baldwin, C.T., Steinberg M.H., Klings, E.S. A genome-wide association study of total bilirubin and cholelithiasis risk in sickle cell anemia 2011 (submitted).

Milton, J.N., Sebastiani, P., Zhang, Y., Nouraie, M., Lee, J. Baldwin, C.T., Zhao, X., Xiong, Z., Zeng, Q., Kato, G.J., Goldsmith, J.C., Taylor, J.G., Gordeuk, V.R., Machado, R.F., Steinberg, M.H., Gladwin, M.T. Clinical and genetic variability of red blood cell hemolysis in sickle cell disease. Blood ASH 2011.

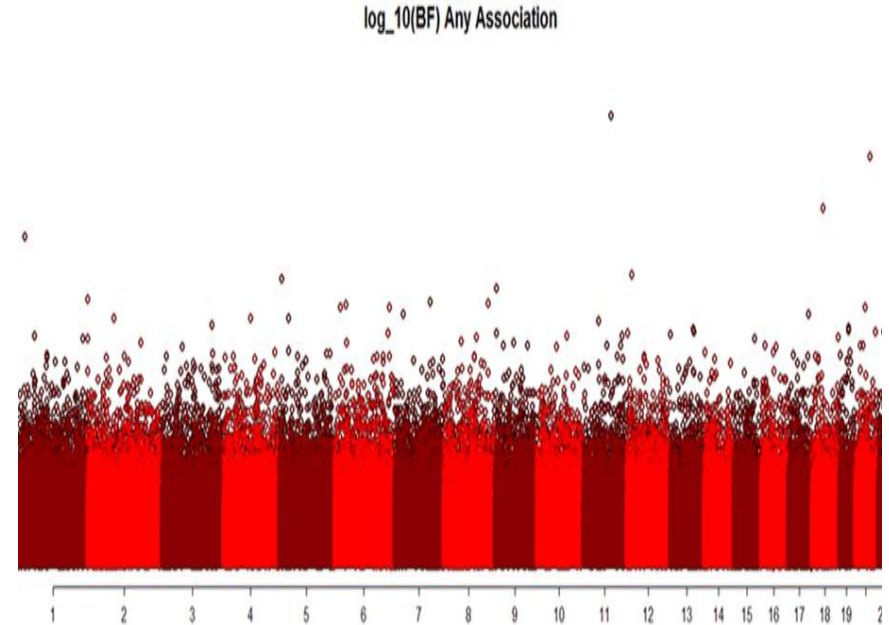
Solovieff, N., Hartley, S.W., Baldwin, C.T., Klings, E.S., Gladwin, M.T., Taylor, J.G.IV, Kato, G.J., Farrer, L.A., Steinberg, M.H., Sebastiani P. Ancestry of African Americans with sickle cell disease. 2011 Blood Cells Mol and Dis 47: 41, 2011.

SickleGen GWAS Studies (others)

TRV: *CSMD1*



Severity score: *KCNK6, TNKS*



Bae, H., Baldwin, C.T., Gladwin, M.T., Ashley-Koch, A.E., Garrett, M., Soldano, K., G. Taylor, J.G., Kato, G.J., Telen, M.T., Sebastiani, P., Steinberg, M.H., Klings, E.S. An elevated tricuspid regurgitant jet velocity in sickle cell disease is associated with polymorphisms in genes impacting innate immunity . Blood ASH 2011.

Sebastiani, P., Timofeev, N., Hartley, S.W., Milton, J.N., Riva, A., Dworkis, D.A., Klings, E.S., Garrett, M.E., Telen, M.J., Ashley-Koch, A., Baldwin, C.T., Steinberg, M.H. Genetic modifiers of the severity of sickle cell anemia identified through a genome-wide association study. Am. J. Hematol 85: 29, 2010.

SickleGen: in progress and in the future

Other subphenotypes for GWAS

Heritability of traits

Hematologic parameters

Pain

Acute chest syndrome

Blood pressure

Nephropathy

HbF response to HU

Meta-analysis of HbF

HbF in Saudi Arabs

Next-gen sequencing and RNA-seq

Saudi Arabs

"Discordant" sib-pairs

QTL-specific iPS cells