FHS 2019 Omics Retreat

2:00PM – 2:05PM	**Welcome**	**Vasan S. Ramachandran, MD** Director and Principal Investigator Framingham Heart Study
2:05PM – 2:40PM	**What data exists?** Overview of FHS Omics resource by exam and generation	**Nancy Heard-Costa, PhD** Manager, Genetic Data Management Group Framingham Heart Study
2:40PM – 3:05PM	**What can we do with this data?** Potential Grants for Phenotyping Working Groups	**Vasan S. Ramachandran, MD** Director and Principal Investigator Framingham Heart Study
3:05PM – 3:25PM	**How would you like statisticians to help?** Descriptive and generic kinds of power calculations PIs need	**Martin G. Larson, SD** Senior Biostatistician Framingham Heart Study
3:25PM – 4:00PM	**Open Discussion**	**All**

BOSTON UNIVERSITY

FHS 2019 Omics Retreat
Welcome

Vasan S. Ramachandran, MD, DM, FACC, FAHA
Director and Principal Investigator
Framingham Heart Study

BOSTON UNIVERSITY

NHLBI Trans-Omics for Precision Medicine

FHS 2019 Omics Retreat
Nancy Heard-Costa
Wednesday October 23, 2019

NHLBI Trans-Omics for Precision Medicine

- focuses on heart, lung, blood, and sleep disorders
- sponsored by the National Institutes of Health (NIH)

  National Heart, Lung and Blood Institute (NHLBI)
- complements NIH Precision Medicine Initiative and *All of Us* Research Program
- provide deep WGS and other omics data to pre-existing ‘parent’ studies having large samples of human subjects with rich phenotypic characterization and environmental exposure data

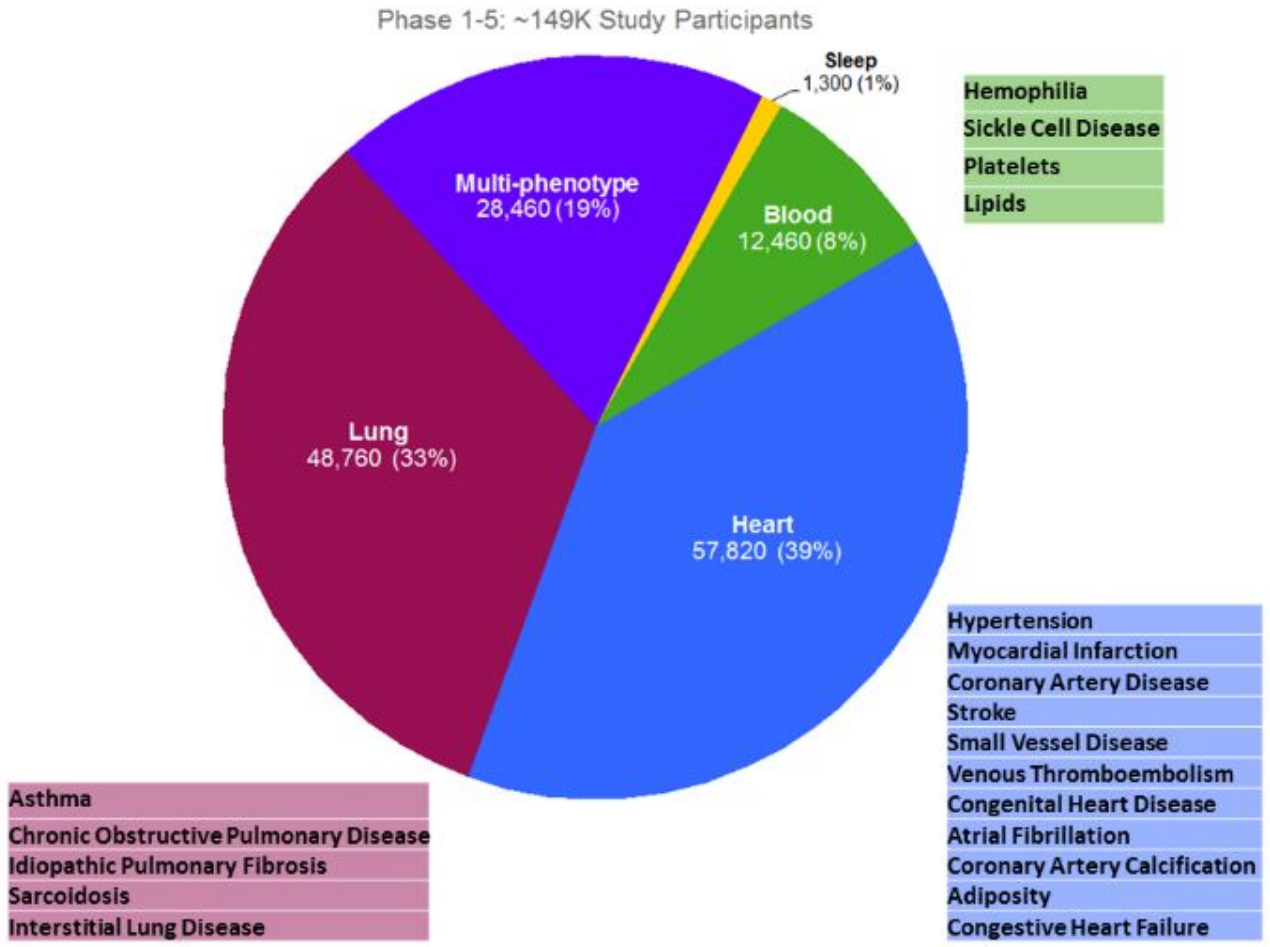
NHLBI Trans-Omics for Precision Medicine

Research Goals

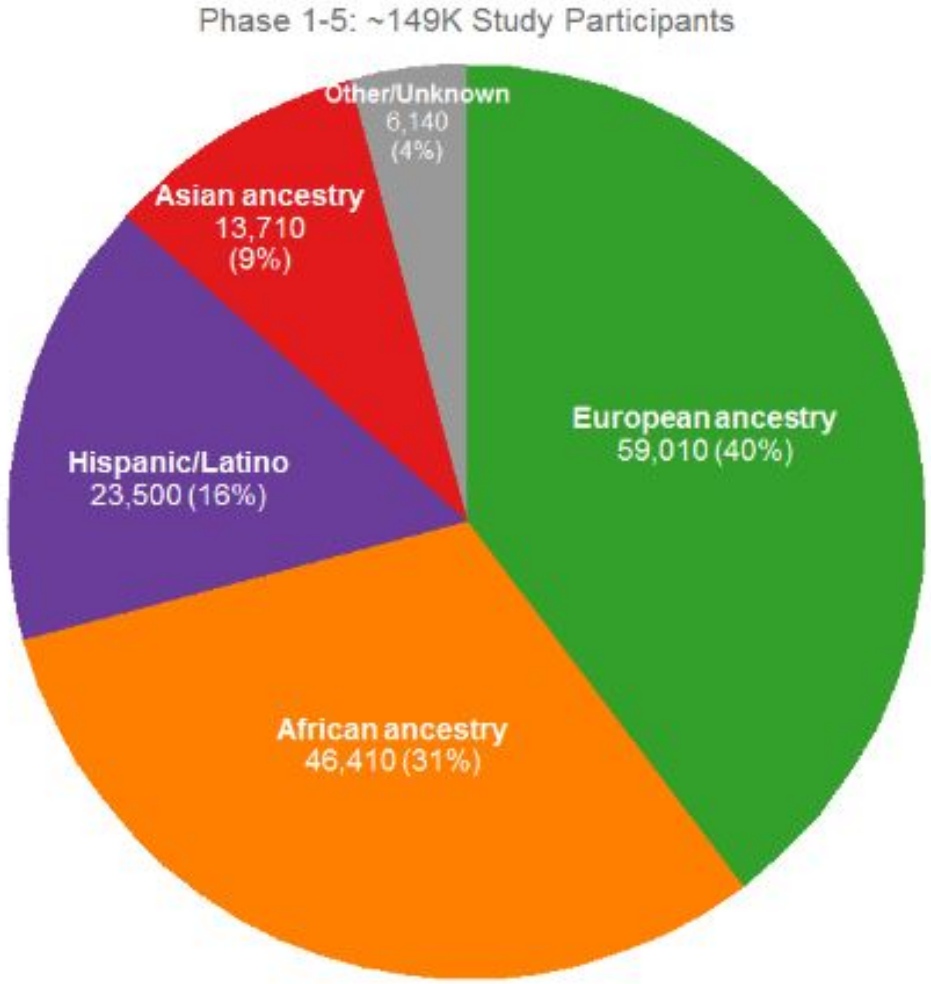
- Biomarkers that increase or decrease the risk of disorders
- Interactions between the environment and genes that affect health
- Potential targets for new treatments
- New ways to disorders or subtypes of these disorders based on molecular signatures
- Targeted ways to develop and test personalized treatments in specific patients
- Advances in precision medicine to predict, prevent, diagnose, and treat heart, lung, blood, and sleep disorders

NHLBI Trans-Omics for Precision Medicine

Sample numbers by phenotype area (N=149k total)



Sample numbers by ancestry/ethnicity (N=149k total)



~149k participants from >80 different studies with varying designs
approximately 60% with substantial non-European ancestry , based on participant annotation and/or study inclusion criteria.

Trans-Omics for Precision Medicine program TOPMed - WGS

- ~ 30X WGS was performed using DNA from blood, PCR-free library construction and Illumina HiSeq X technology.
- Undergo joint genotype calling across all samples available to produce genotype data “freezes.”
- TOPMed data are being made available to the scientific community:
  - genotypes and phenotypes via dbGaP
  - read alignments via the Sequence Read Archive (SRA)
  - variant summary information via the Bravo variant server and dbSNP.

Trans-Omics for Precision Medicine program TOPMed - WGS

Freeze8 includes ~186k samples – among those passing variant QC

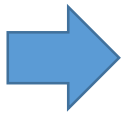
- 811 million single nucleotide variants
- 66 million short insertion/deletion variants
- More information about WGS methods can be found under Sequencing and Data Processing Methods on their website

- https://www.nhlbiwgs.org/

Trans-Omics for Precision Medicine program TOPMed – FHS WGS SAMPLE SELECTION

- Must have blood-based DNA
- Exome Picks*
- Used in WES and WGS studies to pick participants to be sequenced from large family collections
- chose participants in FHS pedigrees *with the best return* for family imputation

SELECTED	
IDTYPE	N
0	377
1	2218
2	95
3	1507
Total	4197



IN FREEZE8	
IDTYPE	N
0	374
1	2210
2	94
3	1499
Total	4177

https://genome.sph.umich.edu/wiki/ExomePicks#Acknowledgements

Trans-Omics for Precision Medicine program TOPMed - X01(Phase5)

- AIM to provide disease treatments tailored to an participant's unique genes and environment
- adding other omic assays *to samples that have been whole-genome sequenced by the program*; these include
  - RNAseq
  - metabolomics
  - epigenomics (DNA methylation)
  - proteomics (Pilot only)

Trans-Omics for Precision Medicine program TOPMed - X01 Metabolomics

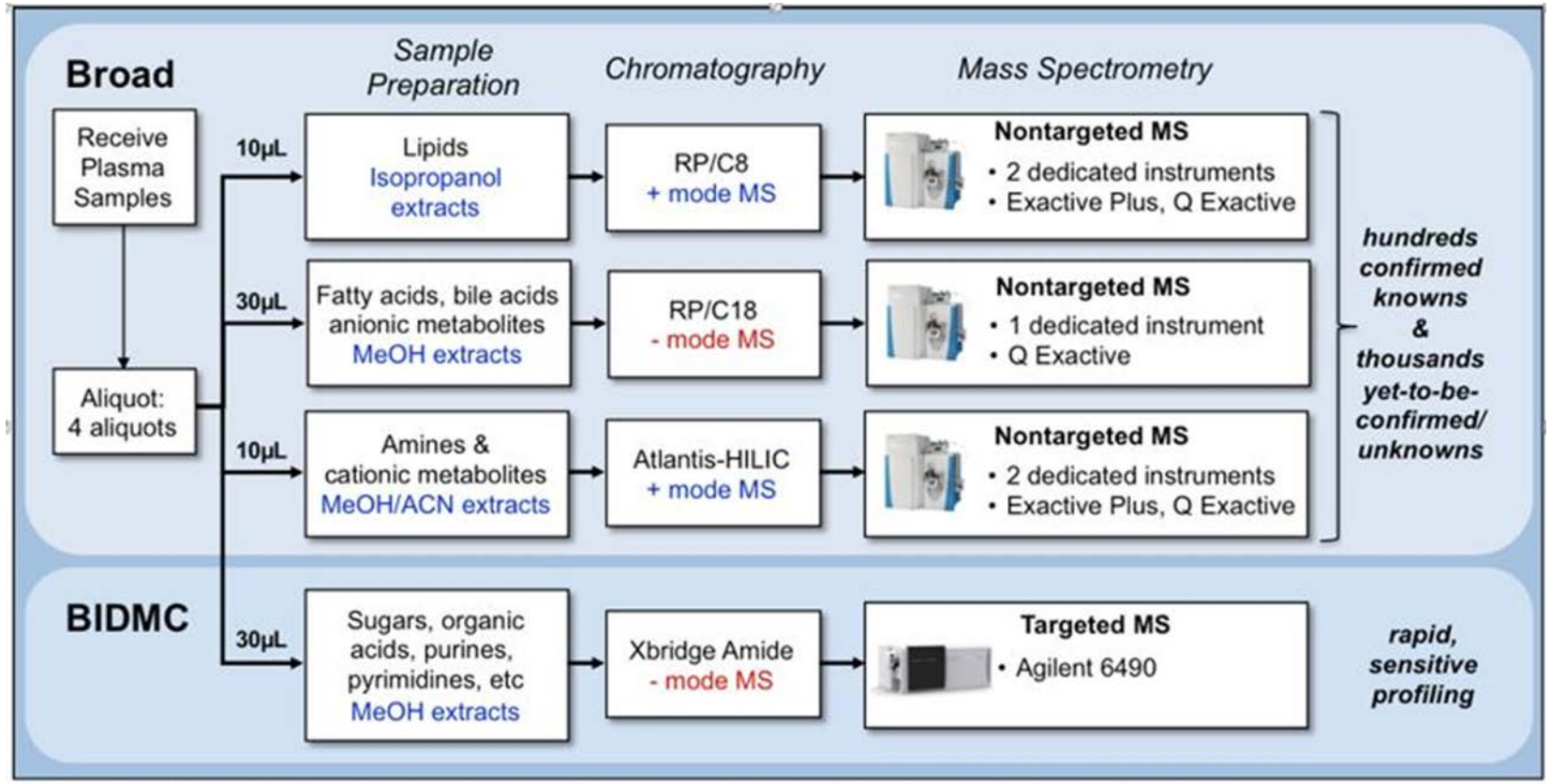


Figure 1. Metabolite profiling workflows and Broad-BIDMC platform integration

Trans-Omics for Precision Medicine program TOPMed - X01 Methylation

- Infinium® MethylationEPIC 850K BeadChip
- over 850,000 methylation sites
- It offers comprehensive, expert-selected coverage, including:
  - 99% of RefSeq genes,
  - 95% of CpG islands,
  - high coverage of enhancer regions, and other content categories.
  - With >90% of the original content from the Infinium HumanMethylation450K covered

Trans-Omics for Precision Medicine program TOPMed - X01 RNA SEQUENCING

- All RNA sequencing performed at the University of Washington Northwest Genomics Center (NWGC), an approved recharge center, directed by Dr. Debbie Nickerson.
- The NWGC carries out all necessary sample processing steps for second-generation sequencing, including RNA quality control/assurance, library construction, targeted, in-solution capture methods, sequencer operation and maintenance, variant calling, data analysis, and IT support
- Same lab and methods as Levy RNASEQ effort

Trans-Omics for Precision Medicine program TOPMed - X01 SELECTION

LIMITED TO PARTICIPANTS WITH EXISTING TOPMED WGS

METABOLOMICS

- EDTA sample
- GEN2 exam 9
- GEN2 exam 8 not present at exam 9
- NOS exam 2
- GEN3 exam 2

METHYLATION

- Blood DNA sample
- GEN2 exam 9
- NOS exam 2
- GEN3 exam 2 with no methylation already & selected for RNA Sequencing

RNA SEQ

- PAXGENE sample
- No existing or expected RNAseq
- GEN2 exam 9
- NOS exam 2
- GEN3 exam 2

Trans-Omics for Precision Medicine program
TOPMED - X01 SELECTION

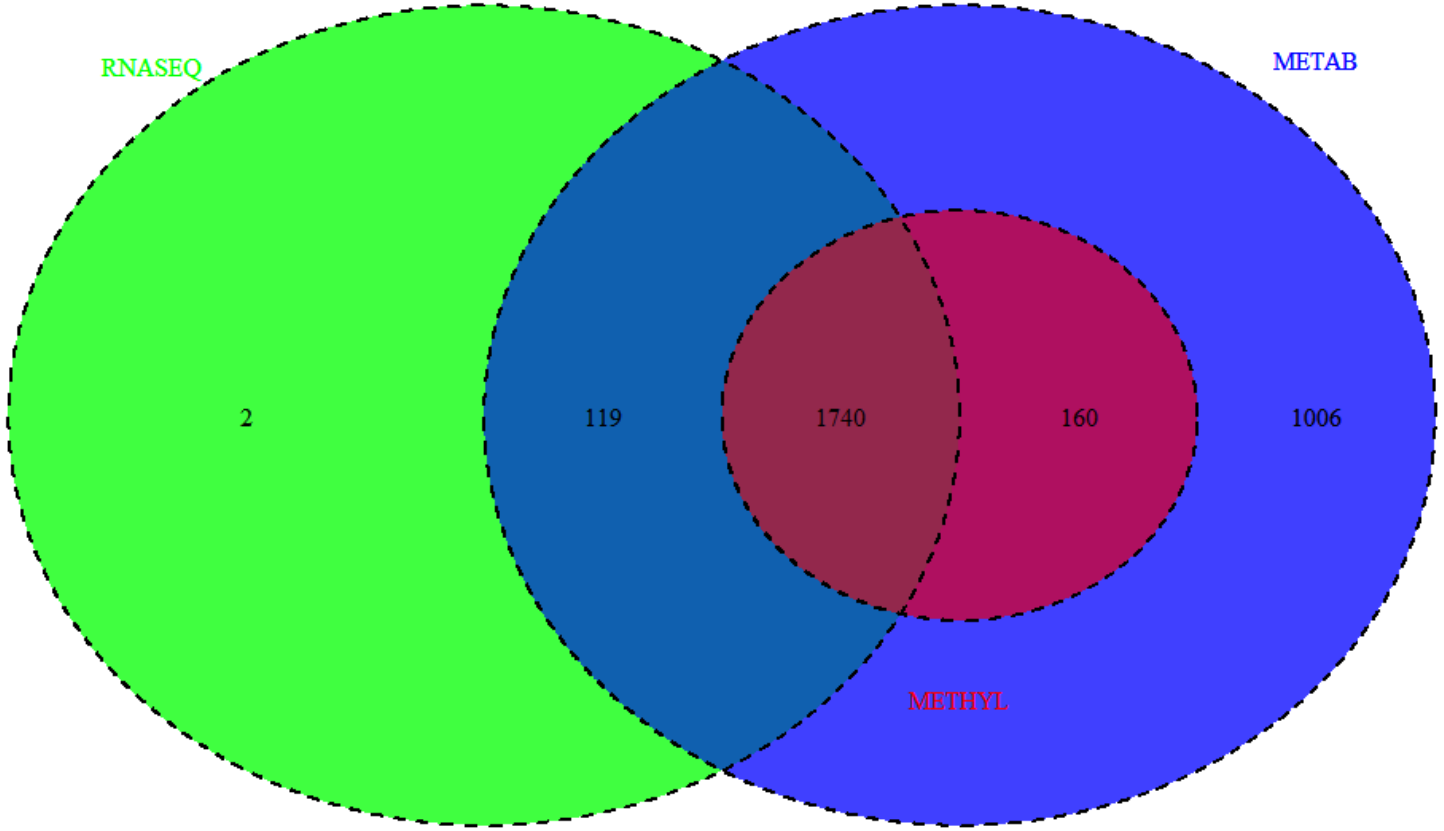
METABOLOMICS SELECT		
IDTYPE	EXAM	N
1	9	1355
1	8	386
2	2	61
3	2	1223
Total		3025

METHYLATION SELECT		
IDTYPE	EXAM	N
1	9	1331
2	2	61
3	2	508
Total		1900

RNASEQ SELECT		
IDTYPE	EXAM	N
1	9	1171
2	2	61
3	2	629
Total		1861

TOPMed - X01 RESOURCE OVERLAP

IDTYPE	EXAM	METAB	METH	RNASEQ
1	9	1355	1331	1171
1	8	386		
2	2	61	61	61
3	2	1223	508	629
Total		3025	1900	1861



FHS LAB:

- Patrice Sutherland
- Chunfeng Kang
- Caitlin Vachon

GENETICS LAB:

- Heather Arruda
- Jessica Rumpf
- Joli Bregu
- Marisol Rodriguez

Trans-Omics for Precision Medicine program TOPMED - X01 ADDITIONAL SEQUENCING

- No prior sequencing (WES/WGS)
  - *exclude* if in TOPMED, CHARGE-S WES, and/or ESP
- Have genotyping
  - *Include* if in SHARE, OMNI CHIP and/or AXIOM CHIP
- DNA extracted from blood
- DNA samples ready to ship
  - On existing blood DNA plates

ADDED WGS SELECT		
IDTYPE	Exam	N
1	8	439
3	2	1922
		2361

Total number with TOPMed WGS = 6538

OTHER FHS OMICS DATA : METABOLOMICS

RESOURCE NAME	N	COHORT/EXAM	DEFINITION
METABOLOMICS			
METABOLO_PHASE5	3025	OFFEX9/8, GEN3/NOS EX2	four complementary metabolite profiling technologies from the Broad Institute and BIDMC
l_mtbgcms	650	Off ex8 & Gen3 ex1	Metabolomics - Risk Factor Study: GC/MS - BMI/Lipids/Glucose Factorial Design
l_mtblcmhi	2067	Off ex5	Central Metabolomics - Hilic -Installments 1&2
l_mtbli	2526	Off ex5	Metabolomics data - Hilic - Installments 1-3
l_mtbllipi	2069	Off ex5	Metabolomics - Lipid Platform - Installment 1&2
l_mtbnegam	998	Gen3 ex1	Negatively Charged Polar Metabolomics - Amide - Installment 1
l_mtbtarg	996	Gen3 ex1	Targeted and Untargeted Metabolomics - HILIC - Installment 1
l_umtbl	386	Off ex6	urine metabolomics

Total number with one or more metabolomics resource(s) available = 5367

OTHER FHS OMICS DATA : METHYLATION AND RNASEQ

RESOURCE NAME	N	COHORT/EXAM	DEFINITION
METHYLATION			
METHYLATION_PHASE5	1900	OFFEX9, GEN3/NOS EX2	Infinium® MethylationEPIC 850K BeadChip
METHYLATION	4151	Off ex8 & Gen3 ex2	Illumina HumanMethylation450 microarray platform
RNA RELATED			
RNASEQ_PHASE5	1861	OFFEX9, GEN3/NOS EX2	Northwest Genomics Center (NWGC) RNA sequencing
RNASEQ_CHRIS	202	OFFEX8	offspring participants with RNA sequence data from Chris O'Donnel
RNASEQ_DAN	1521	GEN3 EX2	RNA samples sent for LEVY; NWGC RNA sequencing

METHYLATION SAMPLES OVERLAP IN GEN2 ONLY - 1202 AT BOTH EXAM 8 AND EXAM 9

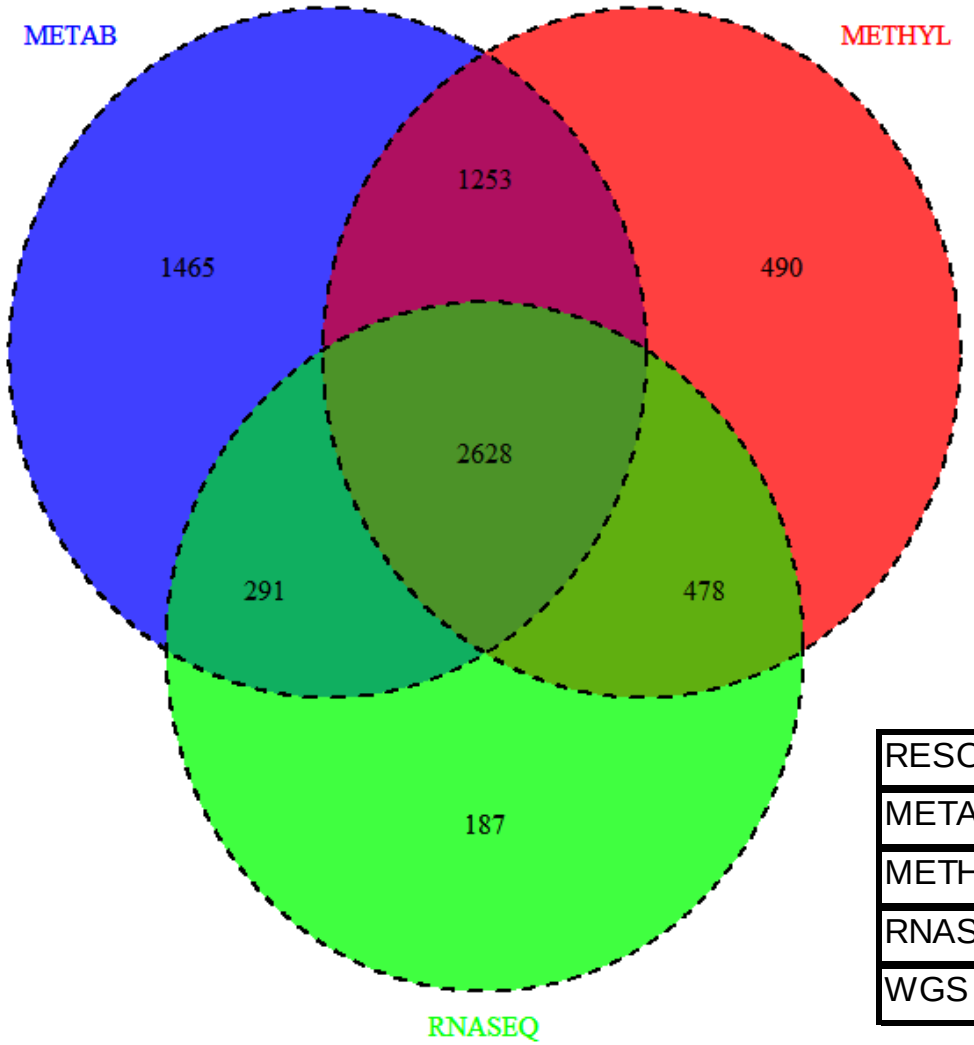
METH_PHASE5			
METH	0	1	TOTAL
0	2364	129	2493
1	1429	1202	2631
TOTAL	3793	1331	5124

Total number with methylation = 4849

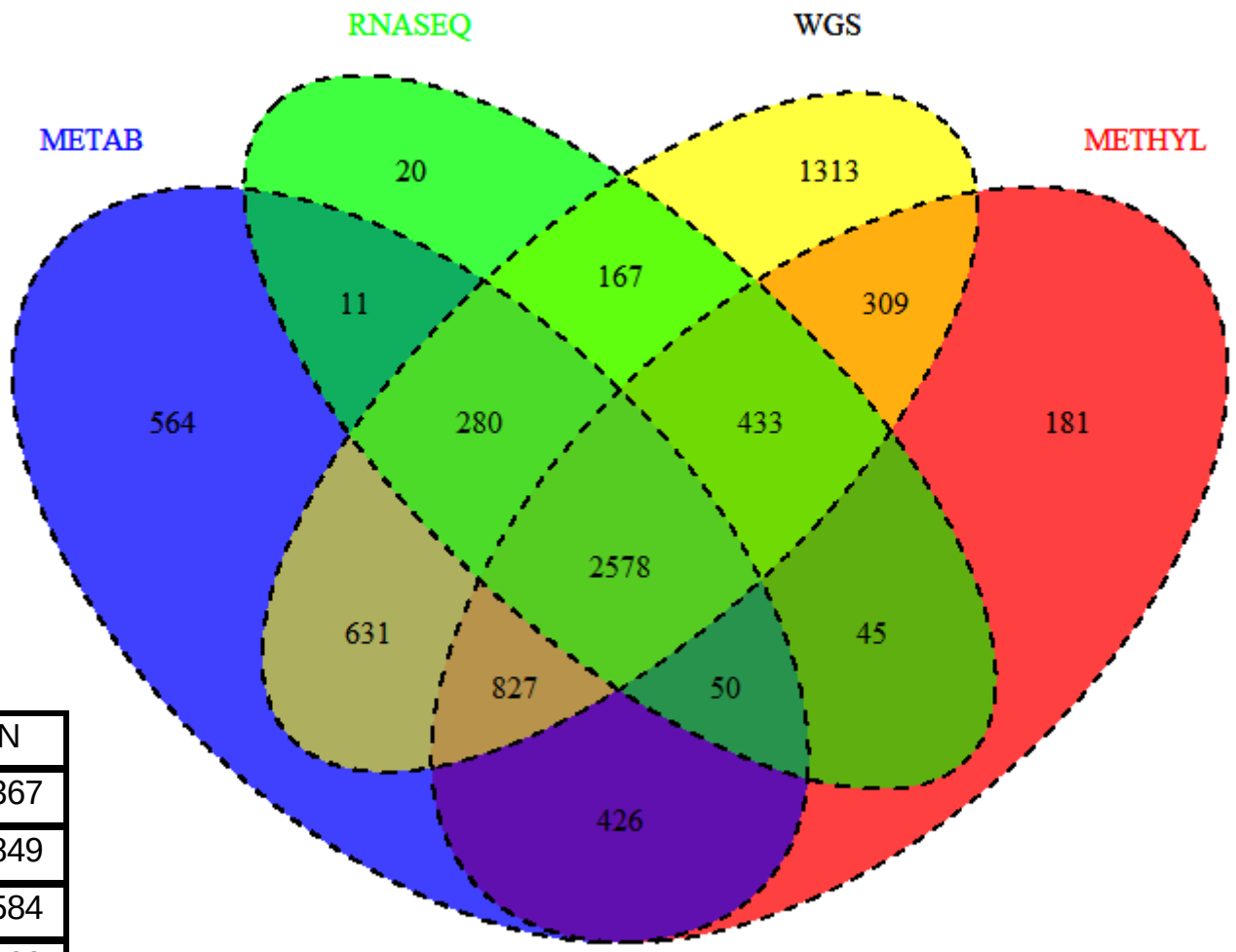
RNA SEQUENCED SAMPLES DO NOT OVERLAP <BY DEMAND>

Total number with RNASEQ = 3584

OTHER FHS OMICS DATA : EXPANDED OVERLAP



RESOURCE	N
METAB	5367
METH	4849
RNASEQ	3584
WGS	6538



ADDITIONAL TYPES OF FHS OMICS DATA

ADDITIONAL TYPES OF FHS OMICS DATA: PROTEOMICS

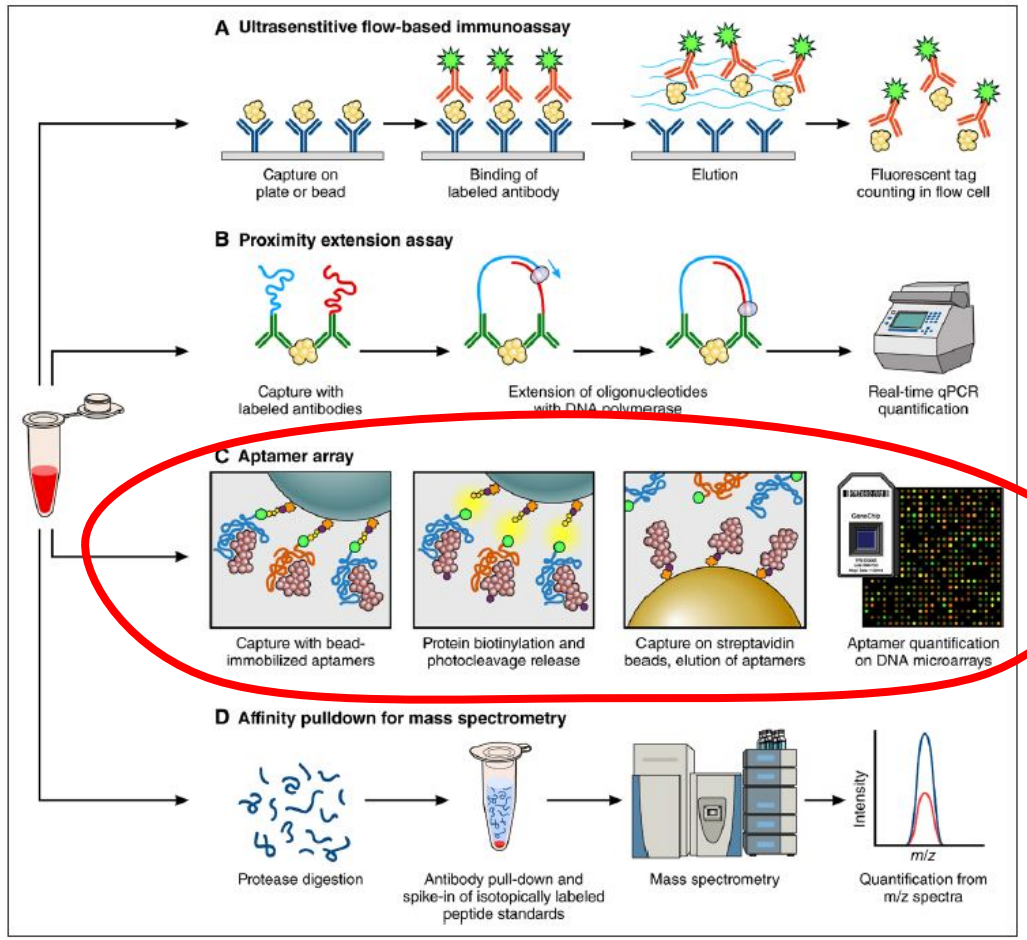


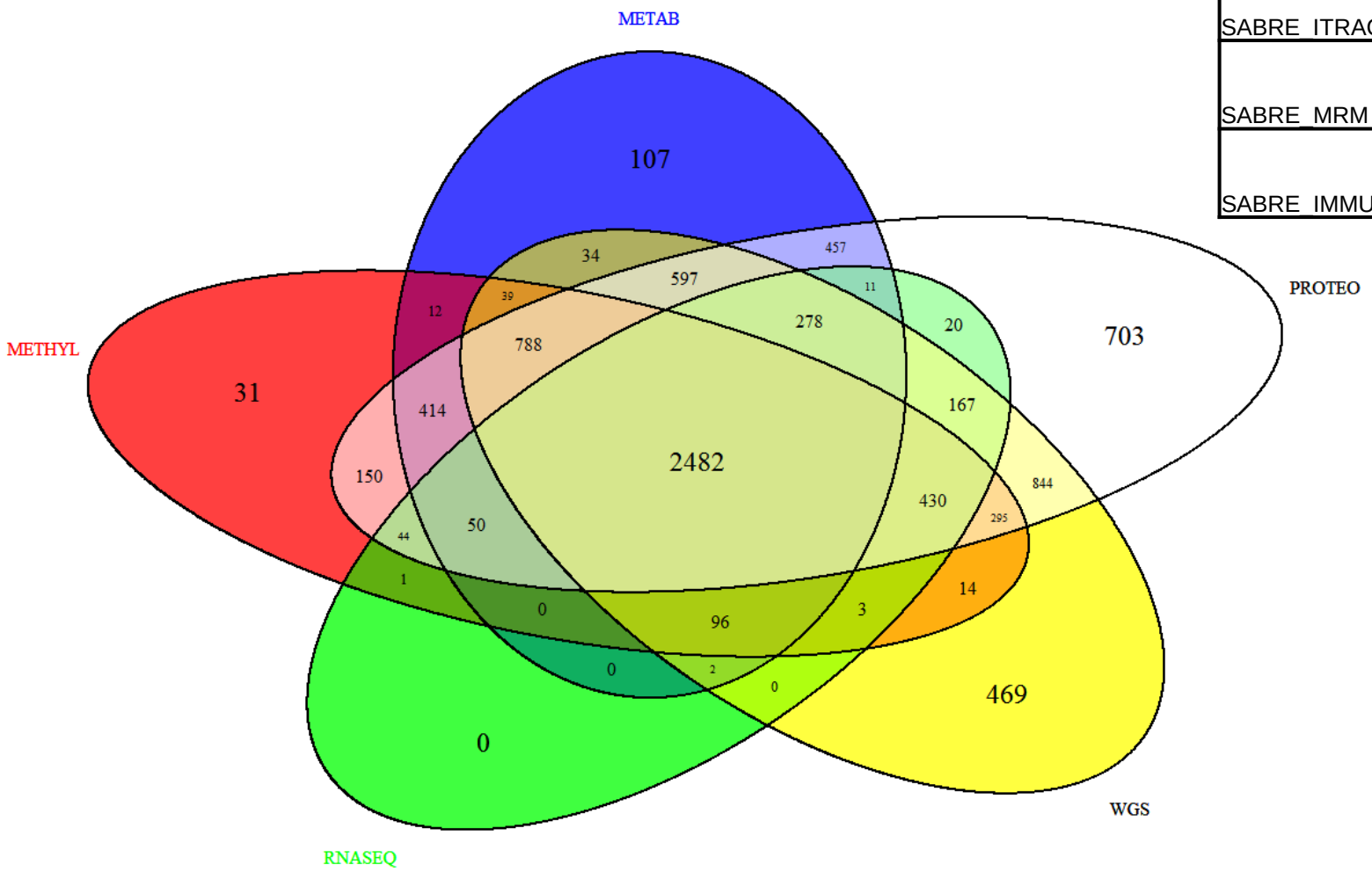
Figure 2. Schematic workflow for emerging affinity proteomic tools.
DNA indicates deoxyribonucleic acid; m/z, mass-to-charge ratio; and qPCR, quantitative polymerase chain reaction.

PROTEOMICS	N	EXAM	DESCRIPTION
PROTEO_TOPMED_PILOT	900	GEN3 EX2	SOMAscan™ proteomic profiling platform (aptamer-based technique)
PROTAPT	1913	Off ex5	SOMAscan™ proteomic profiling platform (aptamer-based technique)
SABRE_ITRAQ	271	Off ex5-ex8	iTRAQ Px data set 135 case/control pairs ;
SABRE_MRM	674	Off ex5-ex8	Targeted MRM Px of 33 targets measured in the CVD study; Multiple reaction monitoring (MRM)
SABRE_IMMUNO	7361	Off ex7 & Gen3 ex1	Immunoassays of ~85 circulating protein biomarkers of atherosclerosis and metabolic syndrome

Total number with one or more proteomics resource = 7730

FHS PROTEOMICS DATA:

Total number with one or more proteomics resource = 7730

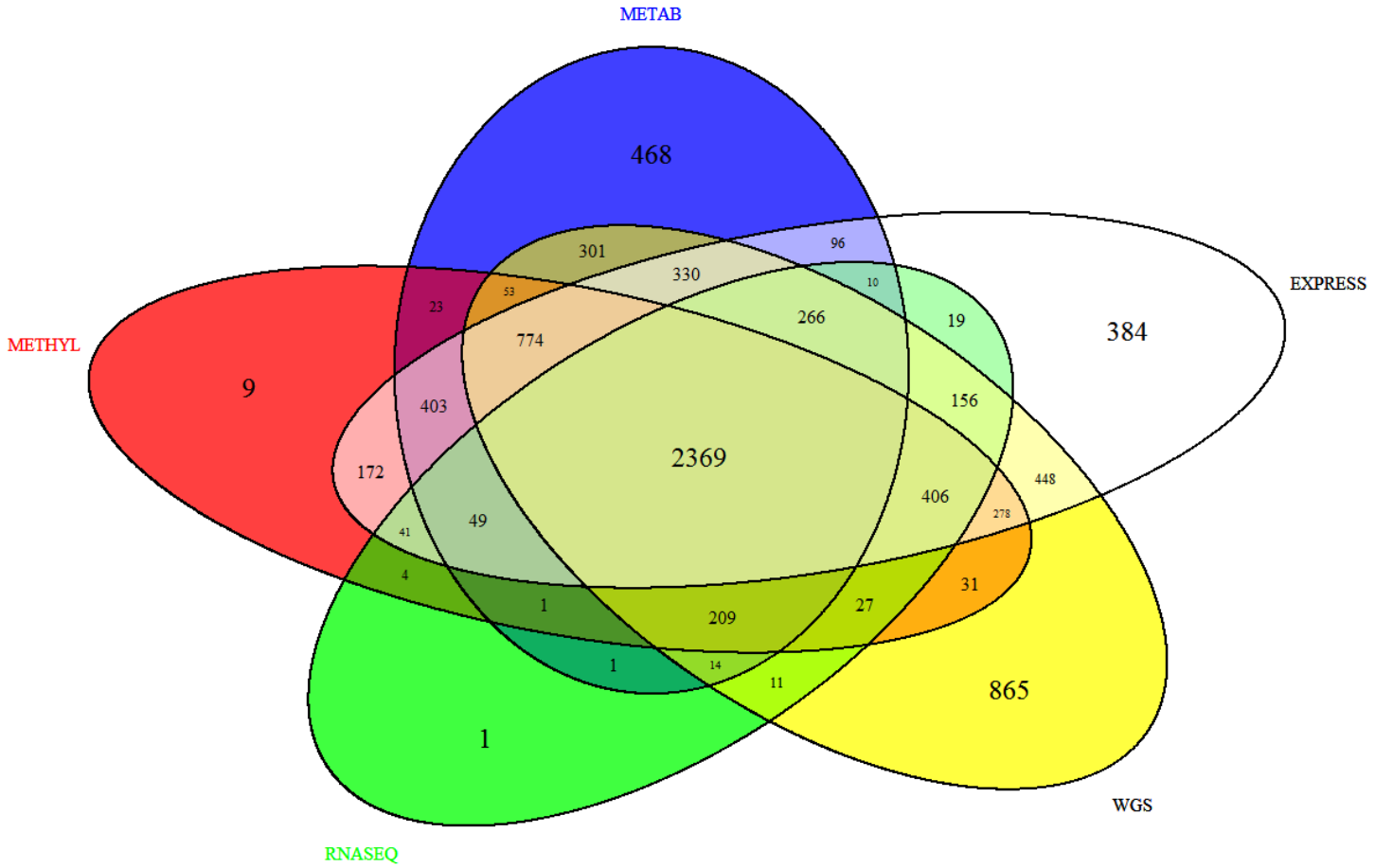


PROTEOMICS	N	EXAM	DESCRIPTION
PROTEO_TOPMED_PILOT	900	GEN3 EX2	SOMAscan™ proteomic profiling platform (aptamer-based technique)
PROTAPT	1913	Off ex5	SOMAscan™ proteomic profiling platform (aptamer-based technique)
SABRE_ITRAQ	271	Off ex5-ex8	iTRAQ Px data set 135 case/control pairs ;
SABRE_MRM	674	Off ex5-ex8	Targeted MRM Px of 33 targets measured in the CVD study; Multiple reaction monitoring (MRM)
SABRE_IMMUNO	7361	Off ex7 & Gen3 ex1	Immunoassays of ~85 circulating protein biomarkers of atherosclerosis and metabolic syndrome

FHS EXPRESSION OMICS DATA

EXPRESSION			
RTPCR	2237	Off ex8 & OMNI1 exam3	RTPCR Gene Expression
EXPRESSION	5626	Off ex8 & Gen3 ex2	gene expression profiling of WBC derived RNA

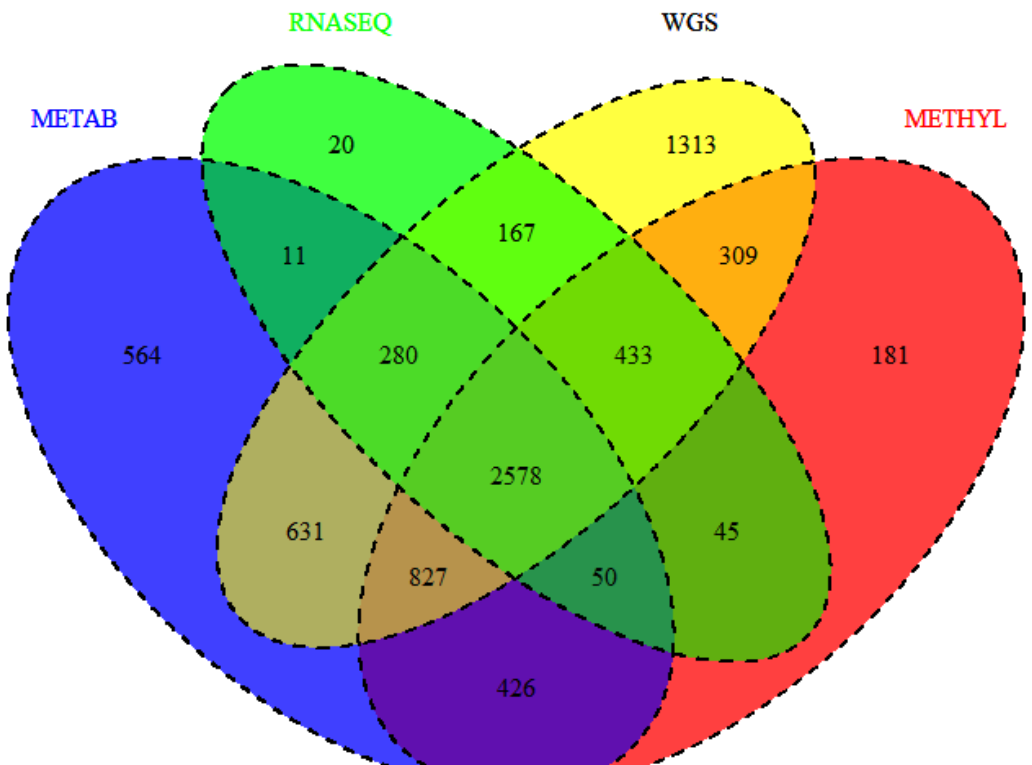
Total number of FHS participants with expression resource = 6201



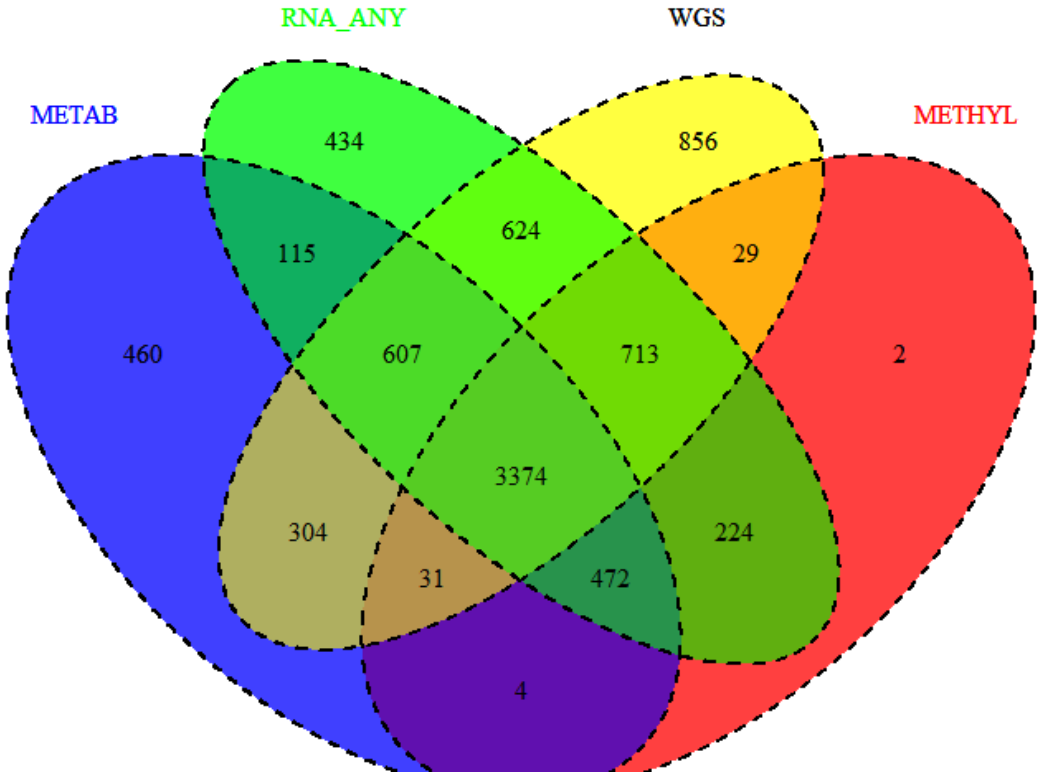
TYPES OF FHS RNA OMICS DATA

RNA RELATED			
RNASEQ_PHASE5	1861	OFFEX9, GEN3/NOS EX2	Northwest Genomics Center (NWGC) RNA sequencing
RNASEQ_CHRIS	202	OFFEX8	offspring participants with RNA sequence data from Chris O'Donnel
RNASEQ_DAN	1521	GEN3 EX2	Northwest Genomics Center (NWGC) RNA sequencing
SABRE_miRNA	5729	Off ex8 & Gen3 ex2	MicroRNA profiling of WBC derived RNA
EXRNA	3071	OFFEX8-9, OMNI1 EX3-4	extra-cellular short RNA

Total number of FHS participants with RNASEQ = 3584



Total number of FHS participants with any RNA resource = 6563



ADDITIONAL FHS OMICS DATA

SEQUENCING	N	COHORT/EXAM	DESCRIPTION
TOPMED_FREEZE8	4177	ALL MAIN COHORTS- VARIED	Eigth freeze of whole genome sequencing produced as part of the TOPMED (dbGaP accession #phs000974)
TOPMED_ADDED	2361	OFFEX8, GEN3EX2	List of participants sent for WGS AFTER original TOPMED samples
CHARGES_WES_FREEZE5	1702	ALL MAIN COHORTS- VARIED	FINAL release of exome sequencing produced as part of CHARGE-S (dbGaP accession #phs000651)
CHARGES_WGS_FREEZE3	855	ALL MAIN COHORTS- VARIED	FINAL release of whole genome sequencing produced as part of CHARGE-S(dbGaP accession #phs000974)
ESP	464	ALL MAIN COHORTS- VARIED	exome sequence data produced as part of NHLBI's GO-ESP project (dbGaP accession #phs000401)
GWAS/SNPS			
SHARE_DEFAULT	8481	ALL MAIN COHORTS- VARIED	SNPs from Affymetrix 500K mapping array plus Affymetrix 50K supplemental array - pass QC* participants
OMNI5	2473	OFFSPRING ONLY VARIED	SNPs from Illumina's HumanOmni5M-4v1 array designed to target variation down to 1% minor allele frequency
EXOME CHIP	8153	ALL MAIN COHORTS- VARIED	SNPS from Illumina Human Exome BeadChips of putative functional exonic variants selected from over 12,000 individual exome and whole-genome sequences through a close collaboration with leading geneticists with the goal of developing an extensive catalog of exome variants.
AXIOM_SNP	845	OMNI1&2 EX1	SNPS from Affymetrix standard Axiom Genome-Wide BioBank array configuration

*exclusion criteria: Per subject callrate < 97%; Per subject heterozygosity ±5 SD away from the mean;
Per subject Mendelian error rate > 2 (among those with call rate > 97%)

The 8481 'SHARE_DEFAULT' participants are included in *population based imputed data* including

- 1000G
- HRC reference panel release 1.1
- TOPMed imputation freeze 5

FHS family-based imputation of TOPMED WGS using Morgan and Gigi is ongoing


THE
BIG PICTURE

Nature of Data	FLAG Name	Cohort					
		Gen 1	Gen 2	Gen 2	Gen 3		
		Original	Offspring	New Offspring Spouse	Generation 3	Omni I	Omni 2
METABOLOMICS	METABOLO_PHASE5		X	X	X		
METABOLOMICS	l_mtbgcms		X		X		
METABOLOMICS	l_mtblcmhi		X				
METABOLOMICS	l_mtbli		X				
METABOLOMICS	l_mtbllipi		X				
METABOLOMICS	l_mtbnegam				X		
METABOLOMICS	l_mtbtarg				X		
METABOLOMICS	l_umtbl		X				
METHYLATION	METHYLATION_PHASE5		X	X	X		
METHYLATION	METHYLATION		X		X		
PROTEOMICS	PROTEO_TOPMED_PILOT				X		
PROTEOMICS	PROTAPT		X				
PROTEOMICS	SABRE_ITRAQ		X				
PROTEOMICS	SABRE_MRM		X				
PROTEOMICS	SABRE_IMMUNO		X		X		
RNA	RNASEQ_PHASE5		X	X	X		
RNA	RNASEQ_CHRIS		X				
RNA	RNASEQ_DAN				X		
RNA	SABRE_miRNA		X		X		
RNA	EXRNA		X			X	
EXPRESSION	RTPCR		X			X	
EXPRESSION	SABRE_EXPRESS		X		X		
DNA SEQUENCE	TOPMED_FREEZE8	X	X	X	X		
DNA SEQUENCE	TOPMED_ADDED		X		X		
DNA SEQUENCE	CHARGES_WES_FREEZE5	X	X		X		
DNA SEQUENCE	CHARGES_WGS_FREEZE3	X	X		X		
DNA SEQUENCE	ESP		X		X		
GWAS DATA	SHARE_DEFAULT	X	X	X	X		
GWAS DATA	OMNI5		X				
GWAS DATA	EXOME CHIP	X	X	X	X		
GWAS DATA	AXIOM_SNP					X	X

Overview of FHS OMICs resources across and within FHS cohorts

Overview of FHS OMICs resources across and within FHS cohorts BY EXAM

		COHORT	OFFSPRING(GEN2)					NOS	GEN3		OMNI1			OMNI2
		MIX	EXAM5	EXAM6	EXAM7	EXAM8	EXAM9	EXAM2	EXAM1	EXAM2	EXAM1	EXAM3	EXAM4	EXAM1
METABOLOMICS	METABOLO_PHASE5					X	X	X		X				
METABOLOMICS	I_mtbgcms					X			X					
METABOLOMICS	I_mtblcmhi		X											
METABOLOMICS	I_mtbli		X											
METABOLOMICS	I_mtbllipi		X											
METABOLOMICS	I_mtbnegam								X					
METABOLOMICS	I_mtbtarg								X					
METABOLOMICS	I_umtbl			X										
METHYLATION	METHYLATION_PHASE5						X	X		X				
METHYLATION	METHYLATION					X				X				
PROTEOMICS	PROTEO_TOPMED_PILOT									X				
PROTEOMICS	PROTAPT		X											
PROTEOMICS	SABRE_ITRAQ		X	X	X	X								
PROTEOMICS	SABRE_MRM		X	X	X	X								
PROTEOMICS	SABRE_IMMUNO				X				X					
RNA	RNASEQ_PHASE5						X	X		X				
RNA	RNASEQ_CHRIS					X								
RNA	RNASEQ_DAN									X				
RNA	SABRE_miRNA					X				X				
RNA	extra-cellular short RNA					X	X					X	X	
EXPRESSION	RTPCR					X						X		
EXPRESSION	SABRE_EXPRESS					X				X				
DNA SEQUENCE	TOPMED_FREEZE8	MIX	MIX					MIX	MIX					
DNA SEQUENCE	TOPMED_ADDED					X				X				
DNA SEQUENCE	CHARGES_WES_FREEZE5	MIX	MIX						MIX					
DNA SEQUENCE	CHARGES_WGS_FREEZE3	MIX	MIX						MIX					
DNA SEQUENCE	ESP		MIX						MIX					
GWAS DATA	SHARE_DEFAULT	MIX	MIX					MIX	MIX					
GWAS DATA	OMNI5		MIX											
GWAS DATA	EXOME CHIP	MIX	MIX					MIX	MIX					
GWAS DATA	AXIOM_SNP										X			X

Nature of Data	FLAG Name	Cohort					
		Gen 1	Gen 2	Gen 2	Gen 3		
		Original	Offspring	New Offspring Spouse	Generation 3	Omni I	Omni 2
METABOLOMICS	METABOLO_PHASE5		1741	61	1223		
METABOLOMICS	l_mtbgcms		316		334		
METABOLOMICS	l_mtblcmhi		2067				
METABOLOMICS	l_mtbli		2526				
METABOLOMICS	l_mtbllipi		2069				
METABOLOMICS	l_mtbnegam				998		
METABOLOMICS	l_mtbtarg				996		
METABOLOMICS	l_umtbl		386				
METHYLATION	METHYLATION_PHASE5		1331	61	508		
METHYLATION	METHYLATION		2631		1520		
PROTEOMICS	PROTEO_TOPMED_PILOT				900		
PROTEOMICS	PROTAPT		1913				
PROTEOMICS	SABRE_ITRAQ		271				
PROTEOMICS	SABRE_MRM		674				
PROTEOMICS	SABRE_IMMUNO		3295		4066		
RNA	RNASEQ_PHASE5		1171	61	629		
RNA	RNASEQ_CHRIS		202				
RNA	RNASEQ_DAN				1521		
RNA	SABRE_miRNA		2495		3234		
RNA	EXRNA		2763			308	
EXPRESSION	RTPCR		1952			285	
EXPRESSION	SABRE_EXPRESS		2446		3180		
DNA SEQUENCE	TOPMED_FREEZE8	374	2210	94	1499		
DNA SEQUENCE	TOPMED_ADDED		439		1922		
DNA SEQUENCE	CHARGES_WES_FREEZE5	49	1537		116		
DNA SEQUENCE	CHARGES_WGS_FREEZE3	13	719		123		
DNA SEQUENCE	ESP		291		173		
GWAS DATA	SHARE_DEFAULT	954	3565	97	3865		
GWAS DATA	OMNI5		2473				
GWAS DATA	EXOME CHIP	655	3380	99	4019		

Participant **numbers** in FHS OMICs resources across and within FHS cohorts

		COHORT	OFFSPRING					NOS	GEN3		OMNI1			OMNI2
		MIX	EXAM5	EXAM6	EXAM7	EXAM8	EXAM9	EXAM2	EXAM1	EXAM2	EXAM1	EXAM3	EXAM4	EXAM1
METABOLO_PHASE	METABOLO_PHASE5					386	1355	61		1223				
METABOLOMICS	l_mtbgcms					316			334					
METABOLOMICS	l_mtblcmhi		2067											
METABOLOMICS	l_mtbli		2526											
METABOLOMICS	l_mtbllipi		2069											
METABOLOMICS	l_mtbnegam								998					
METABOLOMICS	l_mtbtarg								996					
METABOLOMICS	l_umtbl			386										
METHYLATION	METHYLATION_PHASE5						1331	61		508				
METHYLATION	METHYLATION					2631				1520				
PROTEOMICS	PROTEO_TOPMED_PILOT									900				
PROTEOMICS	PROTAPT		1913											
PROTEOMICS	SABRE_ITRAQ		68	54	111	38								
PROTEOMICS	SABRE_MRM		170	132	299	73								
PROTEOMICS	SABRE_IMMUNO				3295				4066					
RNA	RNASEQ_PHASE5						1171	61		629				
RNA	RNASEQ_CHRIS					202								
RNA	RNASEQ_DAN									1521				
RNA	SABRE_miRNA					2495				3234				
RNA	extra-cellular short RNA					2763	1940					263	242	
EXPRESSION	RTPCR					1952						285		
EXPRESSION	SABRE_EXPRESS					2446				3180				
DNA SEQUENCE	TOPMED_FREEZE8	374	2210*					94	173*					
DNA SEQUENCE	TOPMED_ADDED					439				1922				
DNA SEQUENCE	CHARGES_WES_FREEZE5	49	1537*						3180*					
DNA SEQUENCE	CHARGES_WGS_FREEZE3	13	719*						116*					
DNA SEQUENCE	ESP		291*						123*					
GWAS DATA	SHARE_DEFAULT	954	3565*					97	3865*					
GWAS DATA	OMNI5		2473*											
GWAS DATA	EXOME CHIP	655	3380*					99	4019*					
GWAS DATA	AXIOM_SNP										464			381

OTHER FHS OMICS DATA : RESOURCES

METABOLOMICS	Data set name(s)
l_mtbgcms	l_mtbgcms_2008_m_0788s
l_mtblcmhi	l_mtblcmhi1_ex05_1_0662s, l_mtblcmhi2_ex05_1_0708s
l_mtbli	l_mtbli1_ex05_1_0610s, l_mtbli2_ex05_1_0660s, l_mtbli3_ex05_1_0707s
l_mtbllipi	l_mtbllipi1_ex05_1_0617s, l_mtbllipi2_ex05_1_0661s
l_mtbnegam	l_mtbnegam1_ex01_3_0956s
l_mtbtarg	l_mtbtarg1_ex01_3_0955s
l_umtbl	l_umtbl_ex06_1_0742s

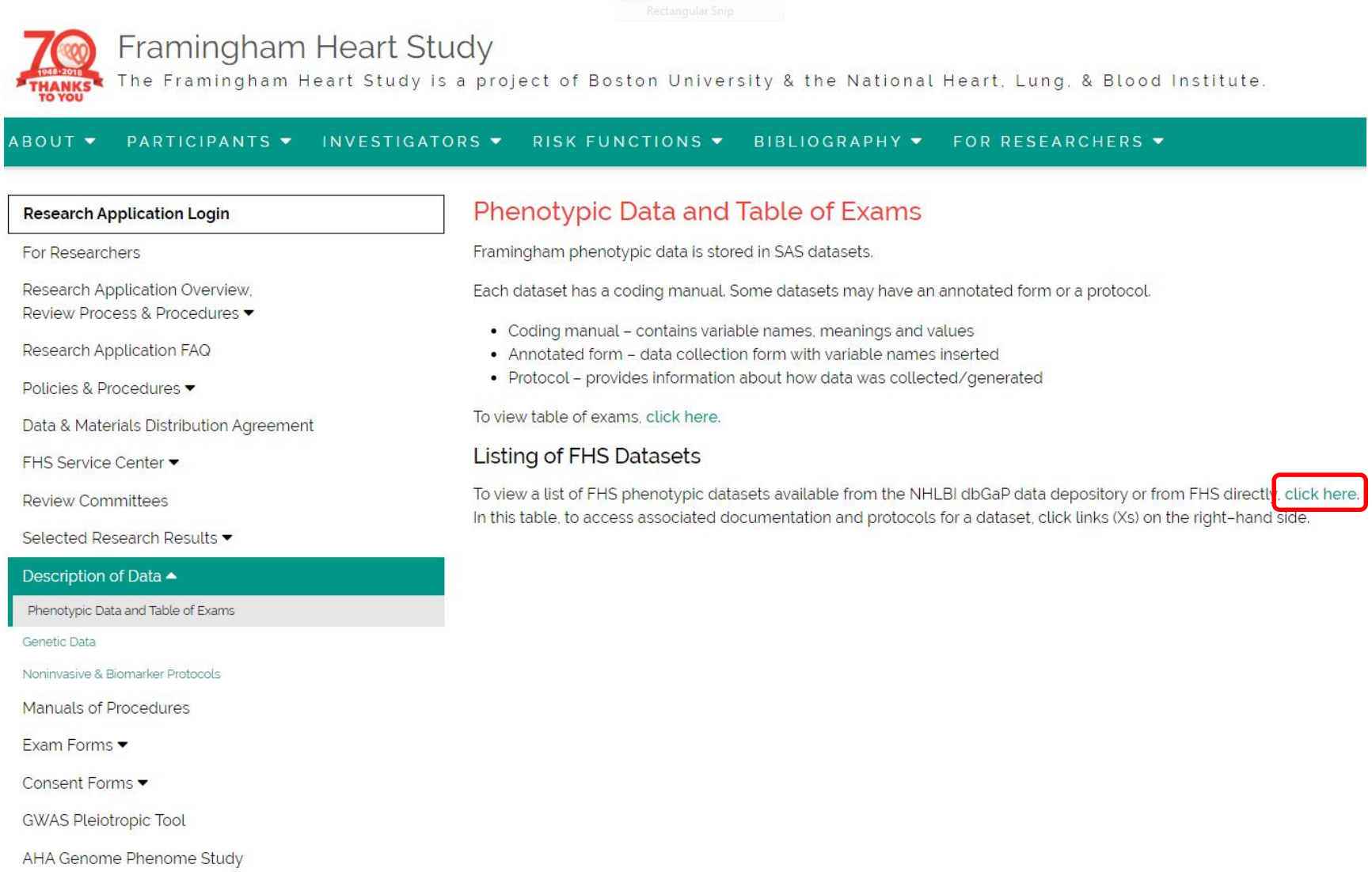
PROTEOMICS	Data set name(s)
PROTEO_TOPMED_PILOT	l_proapt_ex02_3_1117
PROTAPT	l_protapt_ex05_1_0995s
SABRE_ITRAQ	l_protitraq_ex08_1_0736s
SABRE_MRM	l_protmrm_ex08_1_0737s
SABRE_IMMUNO	l_mpimn01_2005_m_0692s, l_mpimn02_2005_m_0693, l_mpimn03_2005_m_0694s, l_mpimn04_2005_m_0757s, l_mpimn05_2005_m_0758s

RNA RELATED	Data set name(s)
SABRE_miRNA	l_mrna_2011_m_0797s

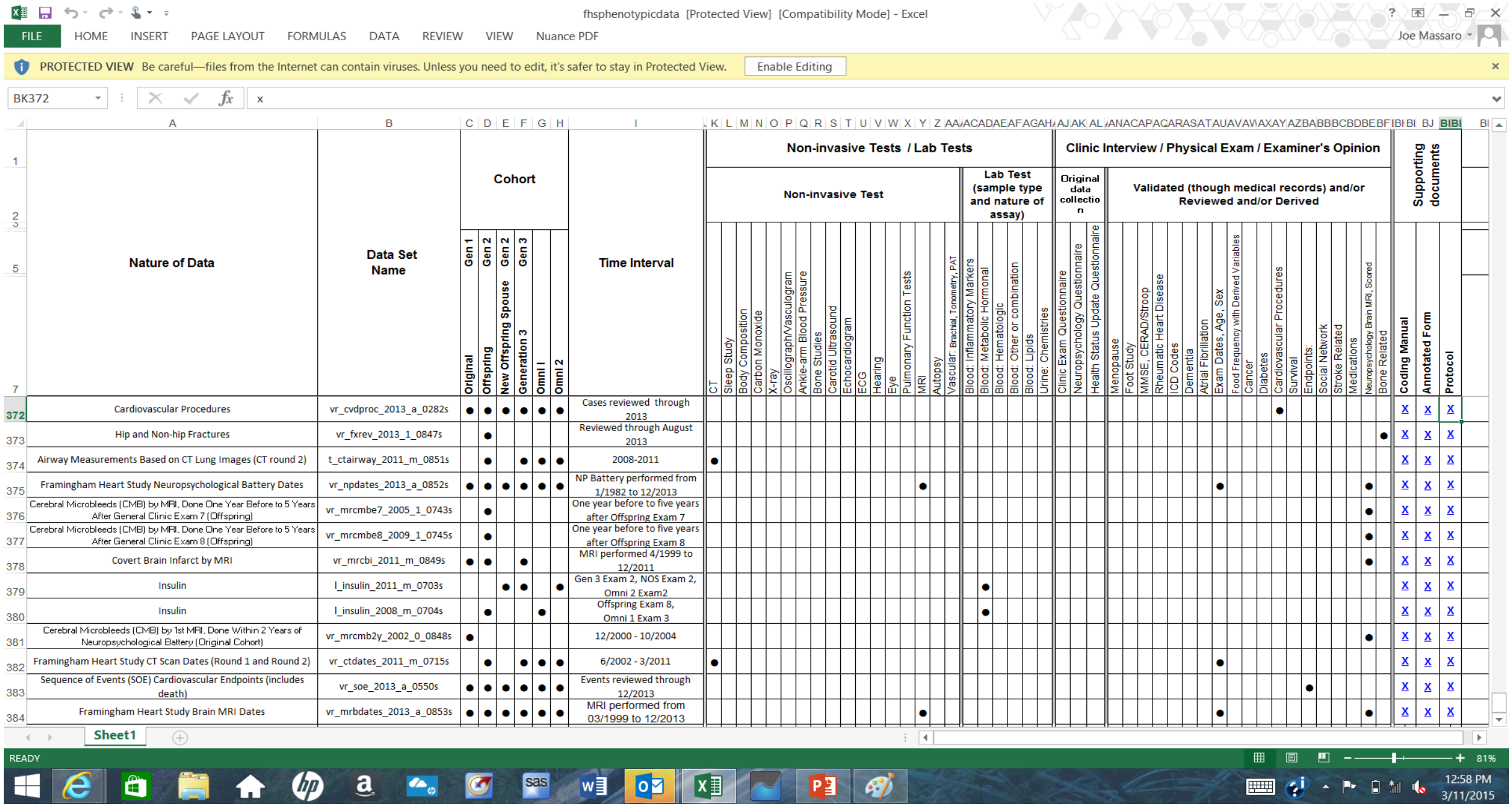
EXPRESSION	Data set name(s)
RTPCR	l_rnatrans_ex08_1_0552s, l_rnatrans2_ex08_1_0619s, l_rnatrns3_ex08_1b_0940s

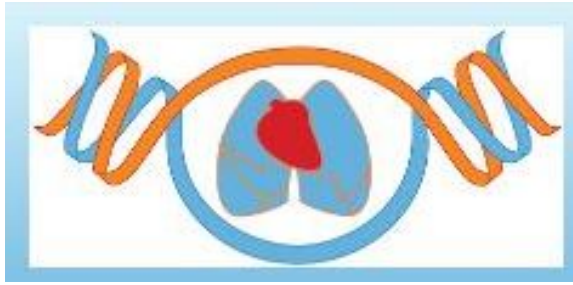
Searching FHS Database

Spreadsheet containing description of phenotypic data collected at
https://www.framinghamheartstudy.org/fhs-for-researchers/phenotypic-data-and-table-of-exams/



Searching FHS Database





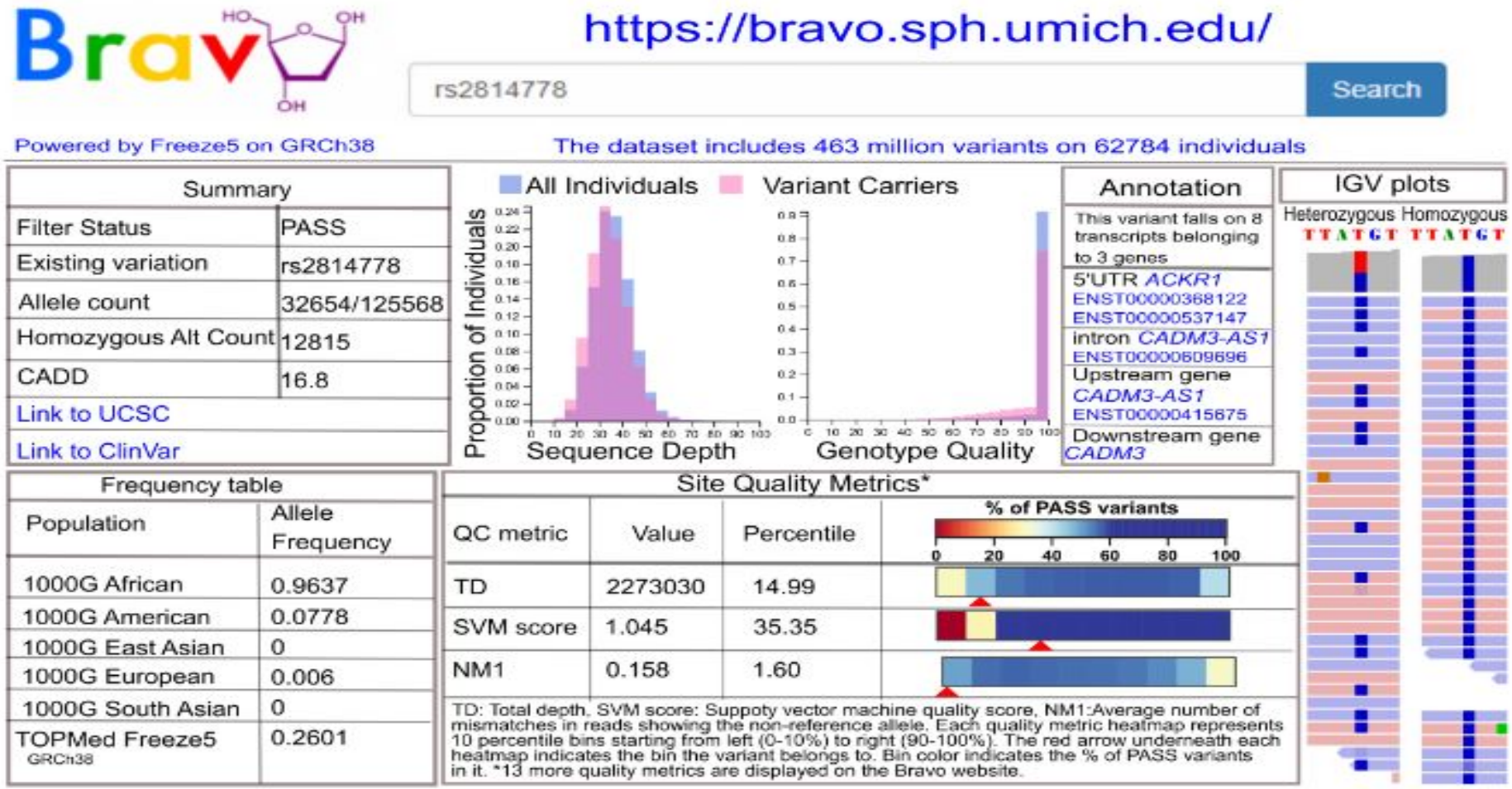
NHLBI Trans-Omics for Precision Medicine

Research Goals

- Biomarkers that increase or decrease the risk of disorders
- Interactions between the environment and genes that affect health
- Potential targets for new treatments
- New ways to disorders or subtypes of these disorders based on molecular signatures
- Targeted ways to develop and test personalized treatments in specific patients
- Advances in precision medicine to predict, prevent, diagnose, and treat heart, lung, blood, and sleep disorders

- ADDED SLIDES

Overview of Bravo variant server resources



This content was adapted from a poster presented at the 2018 American Society of Human Genetics (ASHG) meeting, "Overview of the NHLBI Trans-Omics for Precision Medicine (TOPMed) program: Whole genome sequencing of >100,000 deeply phenotyped individuals" (Poster 3145/T).

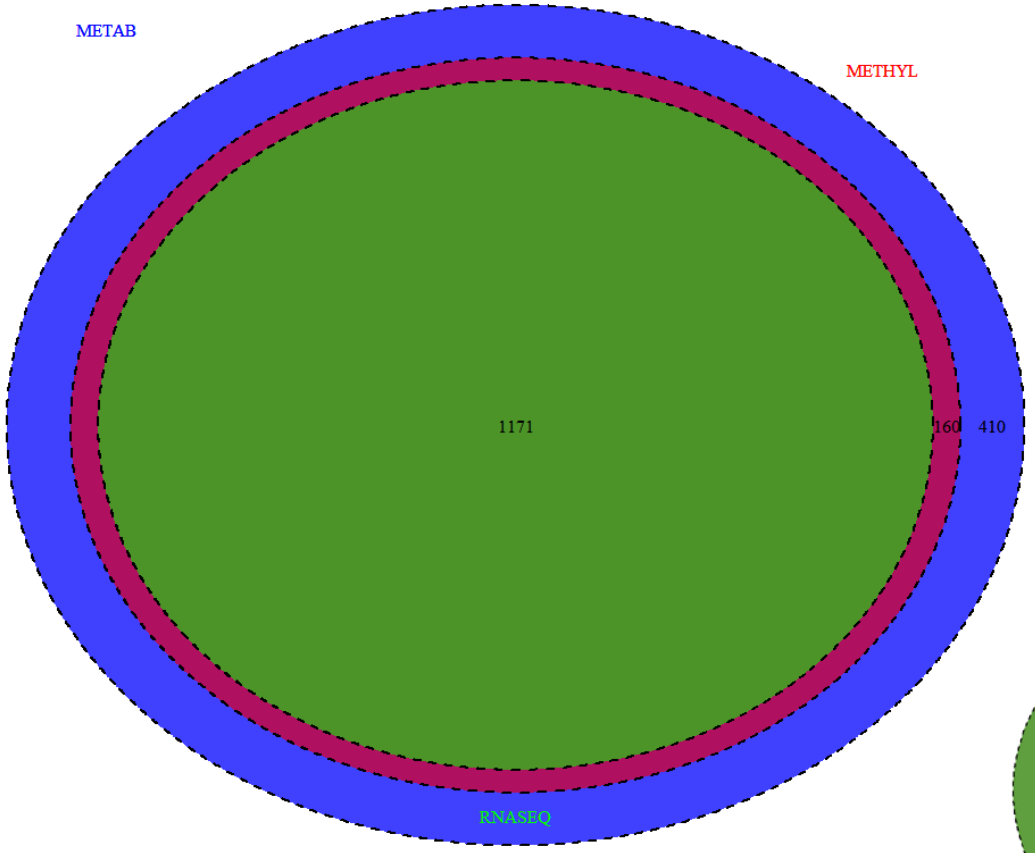
Trans-Omics for Precision Medicine program
TOPMed – EXOME PICKS DETAILS

- loops through sibships, starting at the top of the pedigree and suggests individuals for sequencing as it moves through.
- In pedigrees where DNA samples are available for everyone, it selects every founder (to identify all segregating chromosomes) plus at least one offspring per founder (to determine phase).
- When founder DNA is missing, it selects additional offspring in for each founder couple (if possible) or in sibships internal to the pedigree (if a DNA sample is not available for founder couple offspring, for example)

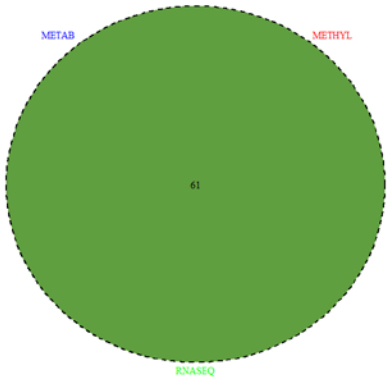
https://genome.sph.umich.edu/wiki/ExomePicks#Acknowledgements

TOPMed – RESOURCE OVERLAP BY COHORT

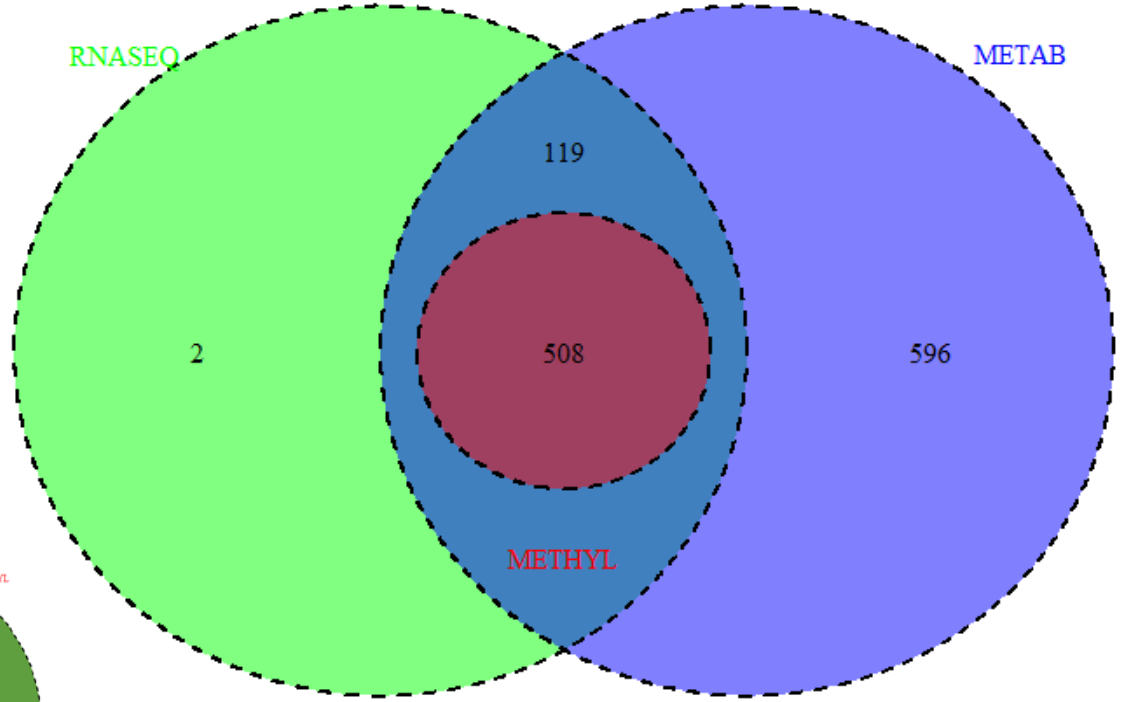
GEN2/OFFSPRING



NOS



GEN3



Variable	N	COHORT/EXAM	DEFINITION
METABOLOMICS			
METABOLO_PHASE5	3025	OFFEX9/8, GEN3/NOS EX2	List of participants sent for metabolomic assay selected from existing FHS TOPMED participants; criteria
l_mtbgcms	650	Off ex8 & Gen3 ex1	Metabolomics - Risk Factor Study: GC/MS - BMI/Lipids/Glucose Factorial Design
l_mtblcmhi	2067	Off ex5	Central Metabolomics - Hilic -Installments 1&2
l_mtbli	2526	Off ex5	Metabolomics data - Hilic - Installments 1-3
l_mtbllipi	2069	Off ex5	Metabolomics - Lipid Platform - Installment 1&2
l_mtbnegam	998	Gen3 ex1	Negatively Charged Polar Metabolomics - Amide - Installment 1
l_mtbtarg	996	Gen3 ex1	Targeted and Untargeted Metabolomics - HILIC - Installment 1
l_umtbl	386	Off ex6	urine metabolomics
METHYLATION			
METHYLATION_PHASE5	1900	OFFEX9, GEN3/NOS EX2	List of participants sent for METHYLATION assay selected from existing FHS TOPMED participants
METHYLATION	4151	Off ex8 & Gen3 ex2	Illumina HumanMethylation450 microarray platform
PROTEOMICS			
PROTEO_TOPMED_PILOT	900	GEN3 EX2	List of participants sent for proteomic assay as part of TOPMED ; not limited to FHS TOPMED participants
PROTAPT	1913	Off ex5	Aptamer Proteomic Profiling: Lab Assay (blood)
SABRE_ITRAQ	271	Off ex5-ex8	iTRAQ Px data set 135 case/control pairs ;
SABRE_MRM	674	Off ex5-ex8	Targeted MRM Px of 33 targets measured in the CVD study; Multiple reaction monitoring (MRM)
SABRE_IMMUNO	7361	Off ex7 & Gen3 ex1	Immunoassays of ~85 circulating protein biomarkers of atherosclerosis and metabolic syndrome
RNA RELATED			
RNASEQ_PHASE5	1861	OFFEX9, GEN3/NOS EX2	List of participants sent for RNA sequencing assay selected from existing FHS TOPMED participants
RNASEQ_CHRIS	202	OFFEX8	offspring participants with RNA sequence data from Chris O'Donnel
RNASEQ_DAN	1521	GEN3 EX2	RNA samples sent for LEVY; same lab and methods as performed in TOPMED
SABRE_miRNA	5729	Off ex8 & Gen3 ex2	MicroRNA profiling of WBC derived RNA
EXRNA	3071	OFFEX8-9, OMNI1 EX3-4	extra-cellular short RNA
EXPRESSION			
RTPCR	2237	Off ex8 & OMNI1 exam3	RTPCR Gene Expression
EXPRESSION	5626	Off ex8 & Gen3 ex2	gene expression profiling of WBC derived RNA
SEQUENCING			
TOPMED_FREEZE8	4177	ALL MAIN COHORTS- VARIED	Eigth freeze of whole genome sequencing produced as part of the TOPMED (dbGaP accession #phs000974)
TOPMED_ADDED	2361	OFFEX8, GEN3EX2	List of participants sent for WGS AFTER original TOPMED samples
CHARGES_WES_FREEZE5	1702	ALL MAIN COHORTS- VARIED	Fifth (and FINAL) release of exome sequencing produced as part of the CHARGE sequencing (CHARGE-S) consortium
CHARGES_WGS_FREEZE3	855	ALL MAIN COHORTS- VARIED	Second (and FINAL) release of whole genome sequencing produced as part of the CHARGE sequencing (CHARGE-S) consortium
ESP	464	ALL MAIN COHORTS- VARIED	exome sequence data produced as part of NHLBI's GO-ESP project (dbGaP accession #phs000401)
GWAS/SNPS			
SHARE_DEFAULT	8481	ALL MAIN COHORTS- VARIED	SNPs from Affymetrix 500K mapping array plus Affymetrix 50K supplemental array - pass QC* participants
OMNI5	2473	OFFSPRING ONLY VARIED	SNPs from Illumina's HumanOmni5M-4v1 array designed to target variation down to 1% minor allele frequency
EXOME CHIP	8153	ALL MAIN COHORTS- VARIED	SNPS from Illumina Human Exome BeadChips of putative functional exonic variants selected from over 12,000 individual exome and whole-genome sequences through a close collaboration with leading geneticists with the goal of developing an extensive catalog of exome variants.
AXIOM_SNP	845	OMNI1&2 EX1	SNPS from Affymetrix standard Axiom Genome-Wide BioBank array configuration

What can we do with this data?
Potential Grants for Phenotyping Working Groups

[Discussion – No Slides]

Vasan S. Ramachandran, MD, DM, FACC, FAHA
Director and Principal Investigator
Framingham Heart Study

BOSTON UNIVERSITY

How would you like statisticians to help?

BOSTON
UNIVERSITY

Statisticians’ help

- Imputation
  - None
  - Single
  - Multiple

Statisticians’ help

- Transformations
  - Logarithmic - ln or log2
  - Standardize - within or across batches
  - Normalize - rank based

Statisticians’ help

- **Power Calculations**
  - **Generic**
  - **Customized**

Statisticians’ help

- Power Calculations – Generic
  - Response variables — Modeled by
    - Continuous — OLS
    - Binary — Logistic
    - Time-to-event — Cox
  - Regressor Variable
    - Continuous
    - Binary

Power Calculations – Generic

Response	Model	Min. Detectable Effect
Continuous	Linear	Partial Correlation
Binary	Logistic	Odds Ratio per 1 SD(X)
Time to Event	Cox	Hazards Ratio per 1 SD(X)

- Regressor, X ~ continuous
- Covariates explain 25% of Var(X)
- Multiple Sample Sizes
- Multiple alpha levels

Power Calculations – linear model

Partial correlations detectable with 80% power
Statistical significance (alpha) per test

Combined Gen2 / Omni1 & Gen3 / Omni2							
	5.00E-02	5.00E-03	5.00E-04	5.00E-05	5.00E-06	5.00E-07	5.00E-08
N=5400	0.038	0.050	0.059	0.067	0.073	0.080	0.085
N=2700	0.054	0.070	0.083	0.094	0.104	0.113	0.121

Power Calculations – logistic model

Odds ratios (per 1 SD of regressor) detectable with 80% power
Statistical significance (alpha) per test

Combined Gen2 / Omni1 & Gen3 / Omni2								
		1 test	10	100	1000	10K	100K	1M
NTotal	Pr(Y=1)	5.00E-02	5.00E-03	5.00E-04	5.00E-05	5.00E-06	5.00E-07	5.00E-08
5400	0.1	1.16	1.21	1.25	1.29	1.33	1.36	1.39
5400	0.2	1.12	1.15	1.19	1.21	1.24	1.26	1.28
5400	0.3	1.10	1.13	1.16	1.18	1.20	1.22	1.24
2700	0.1	1.23	1.31	1.38	1.44	1.49	1.54	1.59
2700	0.2	1.17	1.22	1.27	1.31	1.35	1.39	1.42
2700	0.3	1.15	1.19	1.23	1.27	1.30	1.33	1.36

Power Calculations – Cox model

Hazards ratios (per 1 SD of regressor) detectable with 80% power
Statistical significance (alpha) per test

		Combined Gen2 / Omni1 & Gen3 / Omni2							
			1 test	10	100	1000	10K	100K	1M
		Events	5.00E-02	5.00E-03	5.00E-04	5.00E-05	5.00E-06	5.00E-07	5.00E-08
		200	1.26	1.35	1.42	1.49	1.55	1.61	1.67
		250	1.23	1.31	1.37	1.43	1.48	1.54	1.58
		300	1.21	1.28	1.33	1.39	1.43	1.48	1.52
		350	1.19	1.25	1.31	1.35	1.40	1.44	1.47
		400	1.18	1.23	1.28	1.33	1.37	1.40	1.44