

Standardization of Pharmacogenetic Nomenclature

Lisa Kalman, PhD
Coordinator, GeT-RM
CDC
LKalman@cdc.gov



The Disclaimer:



The findings and conclusions in this report are those of the author and do not necessarily represent the official position of the Centers for Disease Control and Prevention/the Agency for Toxic Substances and Disease Registry.

The use of trademarks in this presentation is for identification only and does not imply endorsement by or other positions of the Department of Health and Human Services or the Centers for Disease Control and Prevention.

Lisa Kalman has no conflicts of interest to report

Reference Materials: Essential for Quality Testing

The Problem

Characterized reference materials are not publicly available for most genetic tests; laboratories often use uncharacterized, non-renewable clinical materials for test validation and QC

Consequences

The quality of genetic tests may be compromised

Genetic Testing Reference Material Program (GeT-RM)

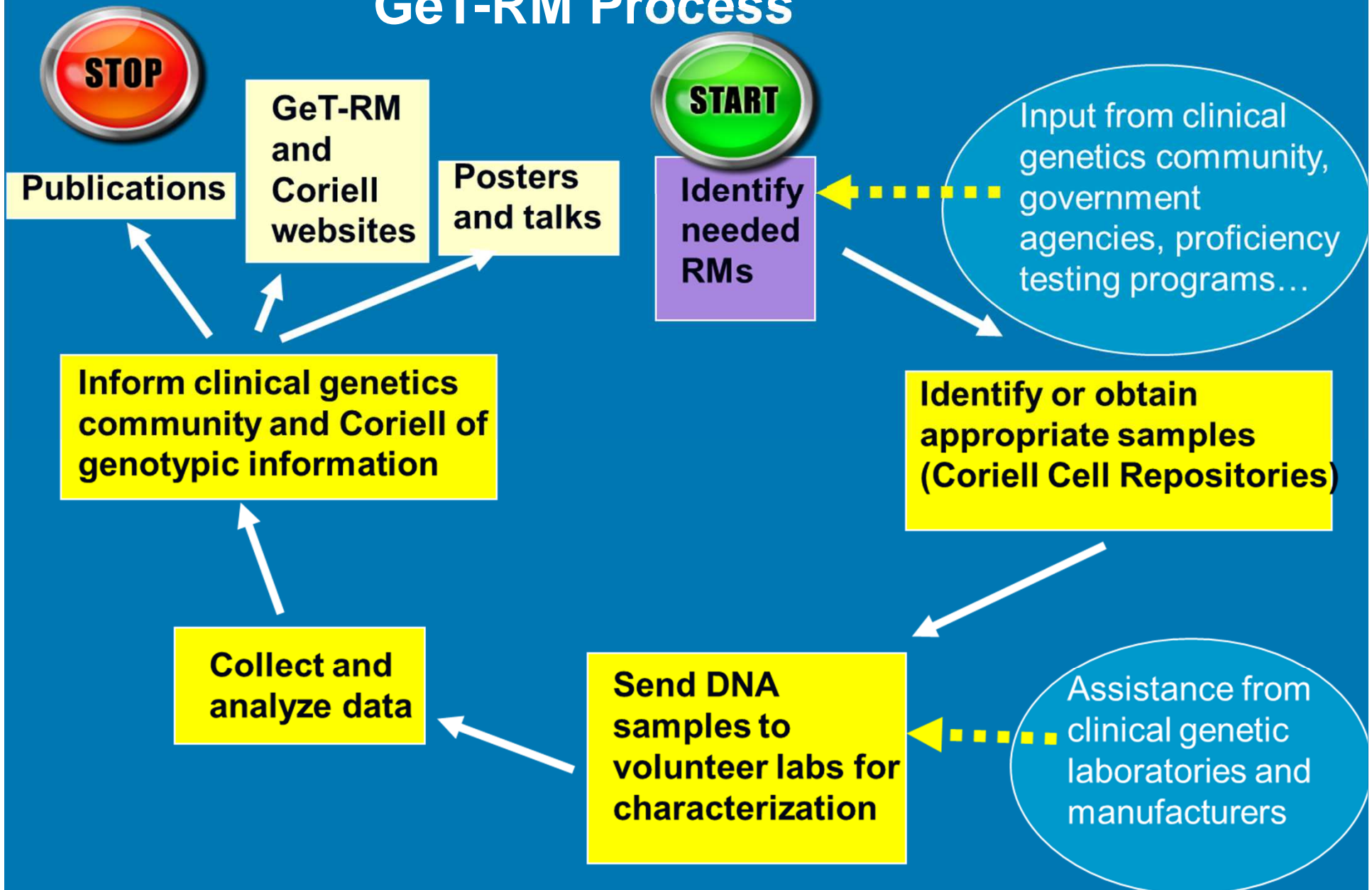
A collaborative CDC-based program to improve the availability of reference materials for genetic testing



Roles of GeT-RM

- Information exchange about RMs for genetic testing
- Monitor RM needs
- Facilitate submission, development and characterization of RMs

GeT-RM Process



GeT-RM Projects

Completed Projects

- Fragile X
- Huntington Disease
- Cystic Fibrosis
- Ashkenazi Jewish Panel
 - 9 disorders including Tay-Sachs and Canavan disease
- BRCA1/2
- MTHFR
- Multiple endocrine neoplasia, Type 2A
- Alpha1-antitrypsin deficiency
- Pharmacogenetics (20 loci)
- Duchenne's muscular dystrophy
- Myotonic dystrophy

Ongoing Projects

- Rett Syndrome
- Next-Generation Sequencing
- Cytogenetics
- Molecular Oncology
- HLA (6 loci)



**Over 400
gDNA RM
characterized
by GeT-RM**

Pharmacogenetics 1

- Involved 13 laboratories (clinical + commercial)
- Assays: laboratory developed tests + many commercial platforms
- Characterized 107 DNA samples for 20 loci (same samples were used in a HLA study)
 - 5 major pharmacogenetic loci (*CYP2D6, CYP2C19, CYP2C9, VKORC1 and UGT1A1*) characterized by multiple labs
 - 9 other pharmacogenetic loci characterized by 1 lab (*CYP4F2, EPHX1, ABCB1, HLAB, KIF6, CYP3A4, CYP3A5, TPMT, DPD*)
 - 6 minor pharmacogenetic loci characterized by 1 or 2 labs (*F5, F2, HFE, MTHFR, AAT, PAI1*)

Characterization of 107 Genomic DNA Reference Materials for CYP2D6, CYP2C19, CYP2C9, VKORC1, and UGT1A1

A GeT-RM and Association for Molecular Pathology Collaborative Project

Victoria M. Pratt,* Barbara Zehnbauer,^{†‡} Jean Amos Wilson,[§] Ruth Baak,[¶] Nikolina Babic,^{||} Maria Bettinotti,* Arlene Buller,** Ken Butz,^{††} Matthew Campbell,^{‡‡} Chris Civalier,^{§§} Abdalla El-Badry,^{¶¶} Daniel H. Farkas,^{¶¶} Elaine Lyon,^{‡‡} Saptarshi Mandal,^{||} Jason McKinney,^{||} Kasinathan Muralidharan,* LeAnne Noll,^{***} Tara Sander,^{***} Junaid Shabbeer,[§] Chingying Smith,[¶] Milhan Telatar,^{†††} Lorraine Toji,^{†††} Anand Vairavan,[¶] Carlos Vance,^{‡‡} Karen E. Weck,^{§§} Alan H.B. Wu,^{§§§} Kiang-Teck J. Yeo,^{||} Markus Zeller,[¶] and Lisa Kalman[†]

From the Quest Diagnostics Nichols Institute,* Chantilly Virginia; the Centers for Disease Control and Prevention,[†] Atlanta, Georgia; Washington University School of Medicine,[‡] St. Louis, Missouri; Berksley HearLab, Inc.,[§] South San Francisco, California; AutoGenomics Inc.,[¶] Vista, California; University of Chicago,^{||} Chicago, Illinois; Quest Diagnostics, Nichols Institute,** San Juan Capistrano, California; ParagonDx, LLC,^{††} Morrisville, North Carolina; ARUP Laboratories,^{‡‡} Salt Lake City, Utah; University of North Carolina,^{§§} Chapel Hill, North Carolina; Sequenom Center for Molecular Medicine,^{¶¶} Grand Rapids, Michigan; Idaho Technology, Inc.,^{||} Salt Lake City, Utah; Children's Hospital of Wisconsin,^{***} Milwaukee, Wisconsin; Specialty Laboratories, Inc.,^{†††} Valencia, California; Coriell Cell Repositories,^{§§§} Camden, New Jersey; and the University of California,^{§§§} San Francisco, California

Pharmacogenetic testing is becoming more common; however, very few quality control and other reference materials that cover alleles commonly included in such assays are currently available. To address these needs, the Centers for Disease Control and Prevention's Genetic Testing Reference Material Coordination Program, in collaboration with members of the pharmacogenetic testing community and the Coriell Cell Repositories, have characterized a panel of 107 genomic DNA reference materials for five loci (CYP2D6, CYP2C19, CYP2C9, VKORC1, and UGT1A1) that are commonly

included in pharmacogenetic testing panels and proficiency testing surveys. Genomic DNA from publicly available cell lines was sent to volunteer laboratories for genotyping. Each sample was tested in three to six laboratories using a variety of commercially available or laboratory-developed platforms. The results were consistent among laboratories, with differences in allele assignments largely related to the manufacturer's assay design and variable nomenclature, especially for CYP2D6. The alleles included in the assay platforms varied, but most were identified in the set of 107 DNA samples. Nine additional pharmacogenetic loci (CYP4F2, EPHX1, ABCB1, HLAB, KIF6, CYP3A4, CYP3A5, TPMT, and DPD) were also tested. These samples are publicly available from Coriell and will be useful for quality assurance, proficiency testing, test development, and research. (*J Mol Diagn* 2010; 12:835–846; DOI: 10.2353/jmol.2010.100090)

Many laboratories are testing for pharmacogenetic (PGx) markers; common genetic variants that are usually considered only when a patient is likely to be exposed to a

Accepted for publication June 21, 2010.

R.B., A.E.-B., C.S., A.V., and M.Z. are employees of AutoGenomics (manufacturer of several pharmacogenetic assays used in this study); M.B., A.B., and K.M. are employees of Quest Diagnostics Inc.; J.M. is an employee of Idaho Technology (manufacturer of the reagents used to genotype CYP2C9 and VKORC1 loci for this project); L.N. and T.S. received donated reagents from Applied Biosystems Inc.; L.T. is employed by the Coriell Institute for Medical Research, Coriell Cell Repositories, which maintains and distributes these biomaterials to investigators for research, education and use as reference materials; K.W. is a consultant to ParagonDx. V.P. is an employee of Quest Diagnostics and owns stock.

Equipment and reagents used for testing was donated by Luminex Inc. There is no monetary support for this work.

Standard of practice is not being defined by this article, and there may be alternatives.

Address reprint requests to Lisa Kalman, Ph.D., Laboratory Research and Evaluation Branch, Division of Laboratory Science and Standards, Office of Surveillance, Epidemiology, and Laboratory Services, Centers for Disease Control and Prevention, 1600 Clifton Road, Mailstop G23, Atlanta, GA 30333. E-mail: L.Kalman@cdc.gov.

Pratt, VM et al.
J Mol Diag 2010 12:835-846

Pharmacogenetics 2

- Second study to increase number of characterized pharmacogenetic genes/alleles
- Compared a number of available PGx panel tests (231 pharmacogenetic loci, over 2000 SNPs)
- Selected 137 Coriell cell lines that contained variants in as many of the SNPs included in the PGx panels as possible

Pharmacogenetics 2

Study Design:

DNA from 137 **Coriell** cell lines are being characterized in 9 labs with a number of PGx platforms:

- Affymetrix DMET (231 genes, 1931 SNPs)
- Sequenom iPLEX ADME (36 genes)
- Sequenom iPLEX expanded panels (*CYP2D6*, *CYP2C19*, *VKORC1*, *CYP2C9*, *UGT1A1*)
- AutoGenomics (*CYP2D6*, *CYP3A4*, *CYP3A5*, *NAT2*)
- GenMark (*CYP2C19*, *CYP3A4/3A5*, *CYP2C9/VKORC1*)
- Luminex (*CYP2D6*, *CYP2C19*, *CYP2C9/VKORC1*)
- Custom Life Tech Taqman array- (17 genes)
- Targeted NGS assays – 3 labs
- LifeTech Ion Torrent NGS assay- 2 labs

Pharmacogenetics 2

Participating Laboratories:

- Robin Everts- [Agena Biosciences \(Sequenom\)](#)
- Uli Broeckel/Praful Aggarwal- [University of Wisconsin](#)
- Stuart Scott- [Icahn School of Medicine at Mount Sinai](#)
- Ruth Baak- [AutoGenomics](#)
- Vicky Pratt- [Indiana University](#)
- Elaine Mardis/Bob Fulton- [Washington University](#)
- Steven Scherer- [Baylor Medical College](#)
- Debbie Nickerson/Adam Gordon- [University of Washington](#)
- Audrey Papp/Wolfgang Sadee- [Ohio State University](#)

**Nomenclature Issues Identified
During the PGx 2 Study**

Nomenclature Issues

Not all assays test the same SNPs for each gene, which can lead to incorrect assignment of haplotypes

Example: CYP2C19

- *4 (**1A>G**; 99C>T, 80161A>G)
- *4B (-3402C>T; **-806C>T**; **1A>G**; 99C>T; 80161C>T)
- *17 (-3402C>T; **-806C>T**; 99C>T; 80161C>T)

If platform does not test for *4 (c.1A>G), how does one distinguish *4B (no activity) or *17 (increased activity)?

Nomenclature Issues

Laboratories report genotype of same gene using different nomenclature

Example 1: SLC22A2

Sample	Lab 1	Lab 2
PGxT117	*1/*2B,*2A/*3A	WT / WT
PGxT118	*3D/*3D	S270A / S270A
PGxT119	*2B/*2B	WT / WT
PGxT120	*1/*2B,*2A/*3A	WT / WT
PGxT121	*1/*1	WT / WT
PGxT122	*1/*3E,*2A/*3D,*2B/*6;*3A/UNK	WT / S270A
PGxT123	-;*1/UNK,*7/UNK	R400C / K432Q
PGxT124	*1/*7	WT / R400C

Nomenclature Issues

Laboratories report genotype of same gene using different nomenclature

Example 2: VKORC1

Sample	Lab 1	Lab2	Lab3	Lab4	Lab5	Lab6	Lab7
PGxT101	H2/H5 / H2/H5	H2/H2	AA	AA	*2/*2	*2 / *2	AA
PGxT102	*2B / *3; H2/H5 / H7B	H2/H7;H6/UNK,H1/UNK,H3/UNK ,H4/UNK,UNK/UNK	GA	GA	*1/*2	*2 / *3	AG
PGxT103	H2/H5 / H2/H5	H2/H2	AA	AA	*2/*2	*2 / *2	AA
PGxT104	*2/H1 / H7B; *2A / *3	H1/H7;H6/UNK,H3/UNK,H4/UNK ,UNK/UNK	GA	GA	*1/*2	*2 / *3	AG
PGxT105	BHT2RE / BHT2RE	H9/H9	GG	GG	*1/*1	*4 / *4	GG
PGxT106	*2/H1 / *2/H1	H1/H1	AA	AA	*2/*2	*2 / *2	AA

Nomenclature Issues

Laboratories report genotype of same gene using different nomenclature

AND
Different assays look at different SNPs

Example 2: VKORC1

Sample	Lab 1	Lab2	Lab3	Lab4	Lab5	Lab6	Lab7
PGxT101	H2/H5 / H2/H5	H2/H2	AA	AA	*2/*2	*2 / *2	AA
PGxT102	*2B / *3; H2/H5 / H7B	H2/H7;H6/UNK,H1/UNK,H3/UNK ,H4/UNK,UNK/UNK	GA	GA	*1/*2	*2 / *3	AG
PGxT103	H2/H5 / H2/H5	H2/H2	AA	AA	*2/*2	*2 / *2	AA
PGxT104	*2/H1 / H7B; *2A / *3	H1/H7;H6/UNK,H3/UNK,H4/UNK ,UNK/UNK	GA	GA	*1/*2	*2 / *3	AG
PGxT105	BHT2RE / BHT2RE	H9/H9	GG	GG	*1/*1	*4 / *4	GG
PGxT106	*2/H1 / *2/H1	H1/H1	AA	AA	*2/*2	*2 / *2	AA

Problems reporting sequence results: which strand?

- The Affymetrix DMET and other assays report genotypes from the coding strand and do not refer to the Human Reference Genome (e.g. hg19)
- Other sequencing assays report from the Human Reference Genome (not always same as coding strand)

Problems reporting sequence results: reference genotype

Example: rs10276036

- Reference genotype for Affymetrix DMET= AA
- Reference genotype for deep sequencing labs = CC

Where does AA come from? It is based on the most common allele observed in HapMap data.

Problems reporting sequence results: SNPs with non-variant calls

- Sequencing labs only report genotypes that are different from the reference sequence (this is how the VCF is usually structured)
- In order to determine the haplotype, it is important to know which SNPs were tested, but found to be the same as reference!

Problems reporting sequence results: SNPs with non-variant calls

rsSNP	Ge- otype data Assay #					Refere- ce ge- otype Assay #				
	1	2	3	4	5	1	2	3	4	5
RS10276036		AG	CT	CT	CT		AA	CC	CC	CC
RS1045642	CT	CT	AG	AG	AG	CC	CC	AA	AA	AA
RS1128501		GG					GG			
RS1128502		AA					AA			
RS1128503	CT	CT	AG	AG	AG	CC	CC	AA	AA	AA
RS1202183		AA					AA			
RS17064		AA					AA			
RS2032581		AA					AA			
RS2032582	GT	GT	AC	AC	AC	GG	GG	AA	AA	AA

↑ Yellow shading indicates discordant data

Problems reporting sequence results SNPs with non-variant

Assays use different reference sequence

rsSNP	Genotype Assay #					Reference genotype Assay #				
	1	2	3	4	5	1	2	3	4	5
RS10276036		AG	CT	CT	CT		AA	CC	CC	CC
RS1045642	CT	CT	AG	AG	AG	CC	CC	AA	AA	AA
RS1128501		GG					GG			
RS1128502		AA					AA			
RS1128503	CT	CT	AG	AG	AG	CC	CC	AA	AA	AA
RS1202183		AA					AA			
RS17064		AA					AA			
RS2032581		AA					AA			
RS2032582	GT	GT	AC	AC	AC	GG	GG	AA	AA	AA

Problems reporting sequence results: SNPs with non-variant calls

Some assays do not report genotypes of SNPs with non-variant genotypes

Reference genotype
Assay #

rsSNP	1	2	3	4	5	1	2	3	4	5
RS10276036		AG	CT	CT	CT		AA	CC	CC	CC
RS1045642	CT	CT	AG	AG	AG	CC	CC	AA	AA	AA
RS1128501		GG					GG			
RS1128502		AA					AA			
RS1128503	CT	CT	AG	AG	AG	CC	CC	AA	AA	AA
RS1202183		AA					AA			
RS17064		AA					AA			
RS2032581		AA					AA			
RS2032582	GT	GT	AC	AC	AC	GG	GG	AA	AA	AA

Nomenclature?

Next Steps:

Develop standardized nomenclature for pharmacogenetics!

- Assembled group of stakeholders including:
 - Testing laboratories
 - Commercial test developers
 - Database providers (NIH/NCBI and gene specific)
 - Gene specific nomenclature groups
 - Pharmacogenomics Research Network (PGRN)
 - PharmGKB
 - Clinical Pharmacogenetics Implementation Consortium (CPIC)
 - HUGO, HGVS Nomenclature curators

- Have started discussions to identify issues and possible solutions

Possible Solutions???

Problem

- Not all assays test the same SNPs for each gene. Lack of standardization can cause ambiguous haplotype assignments.

Solution?

- Create defined list of SNPs that MUST be assayed for each gene (similar to ACMG list of 23 cystic fibrosis variants)-based on CPIC recommendations?
- Labs must clearly state the genotype (variant or not variant) of each SNP on the “must test” list, or indicate that the SNP was not tested.

Possible Solutions???

Problem

- There are many nomenclatures used to report results of a single gene

Solution?

- Use HGVS nomenclature to describe variant SNPs
- Refer to rsSNP #
- Can also state results in * nomenclature to provide continuity

Possible Solutions???

Problem

- Not all platforms define the reference sequence of the same SNPs in the same way (some use the coding strand, some use the positive strand, some use reference genotype from a different population)

Solution?

- Use the Human Reference Genome (e.g. hg19) and/or LRG as reference sequence

Other issues to consider...

- How can non-variant SNPs (i.e. SNPs with the reference sequence) be represented in VCF? (forced calling?)
- How to deal with sequencing data that detects many new variants that do not have described haplotypes?
- How should databases handle changing nomenclature? (esp. related to *alleles)
- How should haplotypes be represented in databases such as ClinVar?

Standardized haplotype nomenclature

Developed prototype tables of VKORC1, TPMT and CYP2C19

- The rsSNPs that compose each haplotype are shown
- Table can be used to translate between different nomenclature systems
- HGVS for each SNP is provided

Standardized haplotype nomenclature

VKORC1: Official Haplotypes and Nomenclature Version: Sept. 30, 2014

yellow- SNPs that define * haplotypes

Developed by Dr. Johan den Dunnen and Dr. Lisa Kalman based on tables from PharmGKB. Contact Lisa Kalman (Lkalman@cdc.gov) for questions or to provide feedback

Genome (GRCh38, hg38)	NC_000016.10	g.31099660G>A	g.31099180G>T	g.31097388C>G	g.31096368C>T	g.31094233A>C	g.31094032G>A	g.31093557G>A	g.31093188C>G	g.31092475A>
Genome (GRCh37, hg19)	NC_000016.9	g.31110981G>A	g.31110501G>T	g.31108709C>G	g.31107689C>T	g.31105554A>C	g.31105353G>A	g.31104878G>A	g.31104509C>G	g.31103796A>
Reference LRG	?									
Reference Gene (RefSeqGene)	NG_011564.1	g.296C>T	g.776C>A	g.2568G>C	g.3588G>A	g.5723T>G	g.5924C>T	g.6399C>T	g.6768G>C	g.7481T>C
Reference Transcript	NM_206824.1	c.-4931C>T	c.-4451C>A	c.-2659G>C	c.-1639G>A	c.173+324T>G	c.173+525C>T	c.173+1000C>T	c.173+1369G>C	c.174-1133T>C
Reference Protein										
Common name?					-1639G>A		6009C>T	6484C>T		7566C>T
LOCATION in gene		5' exon 1	5' exon 1	5' exon 1	5' exon 1	1i	1i	1i	1i	1i
Haplotype Id	VKORC1	rs7196161	rs17880887	rs17881535	rs9923231	rs2884737	rs17708472	rs9934438	rs8050894	rs2359612
PA165816585	H1 (consistent with *2)	G	G	C	T	A	G	A	G	A
PA165816586	H2 (consistent with *2)	G	G	C	T	C	G	A	G	A
PA165816587	H3 (consistent with *1)	G	G	C	C	A	G	G	G	G
PA165816588	H4	G	G	C	C	A	G	G	C	A
PA165816589	H5 (consistent with *2)	A	G	C	T	C	G	A	G	A
PA165816590	H6 (consistent with *1)	A	G	C	C	A	G	G	C	G
PA165816591	H7 (consistent with *3)	A	G	C	C	A	G	G	C	G
PA165816592	H8 (consistent with *3)	A	T	C	C	A	G	G	C	G
PA165816593	H9 (consistent with *4)	A	T	G	C	A	A	G	C	G

**VKORC1: Official Haplotypes
and Nomenclature
Version: Sept. 30, 2014**

yellow- SNPs
that define *
haplotypes

**Developed by Dr. Johan de
based on tables from Pha
(Lkcalman@cdc.gov) for qu**

Genome (GRCh38, hg38)	NC_000016.10	g.31099660G>A	g.31099180G>T	g.31097388C>G	g.31096368C>T	g.31094233A>C
Genome (GRCh37, hg19)	NC_000016.9	g.31110981G>A	g.31110501G>T	g.31108709C>G	g.31107689C>T	g.31105554A>C
Reference LRG	?					
Reference Gene (RefSeqGene)	NG_011564.1	g.296C>T	g.776C>A	g.2568G>C	g.3588G>A	g.5723T>G
Reference Transcript	NM_206824.1	c.-4931C>T	c.-4451C>A	c.-2659G>C	c.-1639G>A	c.173+324T>G
Reference Protein						
Common name?					-1639G>A	
LOCATION in gene		5' exon 1	5' exon 1	5' exon 1	5' exon 1	1i
Haplotype Id	VKORC1	rs7196161	rs17880887	rs17881535	rs9923231	rs2884737
PA165816585	H1 (consistent with *2)	G	G	C	T	A
PA165816586	H2 (consistent with *2)	G	G	C	T	C
PA165816587	H3 (consistent with *1)	G	G	C	C	A
PA165816588	H4	G	G	C	C	A
PA165816589	H5 (consistent with *2)	A	G	C	T	C
PA165816590	H6 (consistent with *1)	A	G	C	C	A
PA165816591	H7 (consistent with *3)	A	G	C	C	A
PA165816592	H8 (consistent with *3)	A	T	C	C	A
PA165816593	H9 (consistent with *4)	A	T	G	C	A

Next Steps....

- The PGx Nomenclature Workgroup will continue to discuss these issues
- Sub-workgroups will develop models to show how our proposed nomenclature could be used to address problems identified in the nomenclature. We plan to make models using 4 genes: TPMT, CYP2C19, VKORC1 and NAT2
- We are preparing a manuscript that describes the issues with PGx nomenclature and our suggestions for addressing these problems

Lisa Kalman: Lkalman@cdc.gov

For more information please contact Centers for Disease Control and Prevention

1600 Clifton Road NE, Atlanta, GA 30333

Telephone: 1-800-CDC-INFO (232-4636)/TTY: 1-888-232-6348

Visit: www.cdc.gov | Contact CDC at: 1-800-CDC-INFO or www.cdc.gov/info

The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the Centers for Disease Control and Prevention.

Center for Surveillance, Epidemiology, and Laboratory Services
Division of Laboratory Programs, Standards, and Services

