NHGRI U24HG010862

PharmCAT: Current Status and Strategies for the Future

October 7, 2021

PharmCAT

The Pharmacogenomic Clinical Annotation Tool

Latest Release: v1.0.0

Software **Download**	Technical **Docs**	View On **GitHub**

http://pharmcat.org

Pharmacogenomics Clinical Annotation Tool

An active area of genomic medicine implementation at many health care organizations and academic medical centers includes development of decision support and return of results around pharmacogenomics. One of the challenges in implementing pharmacogenomics is the representation of the information in clinical dosing guidelines, including star-allele haplotypes, and extracting these variants and haplotypes from genetic datasets. In a collaboration between the Pharmacogenomics Knowledgebase (PharmGKB) and the former PGRN Statistical Analysis Resource (P-STAR), with input from other groups, we are developing a software tool to extract guideline variants from a genetic dataset (represented as a vcf), interpret the variant alleles, and generate a report with genotype-based prescribing recommendations which can be used to inform treatment decisions. The Clinical Pharmacogenetics Implementation Consortium (CPIC) has established guidelines surrounding gene-drug pairs that can and should lead to treatment modifications based on genetic variants. These guidelines are used for the initial version of PharmCAT, and other sources of PGx information and guidelines will be included in the future.

References:

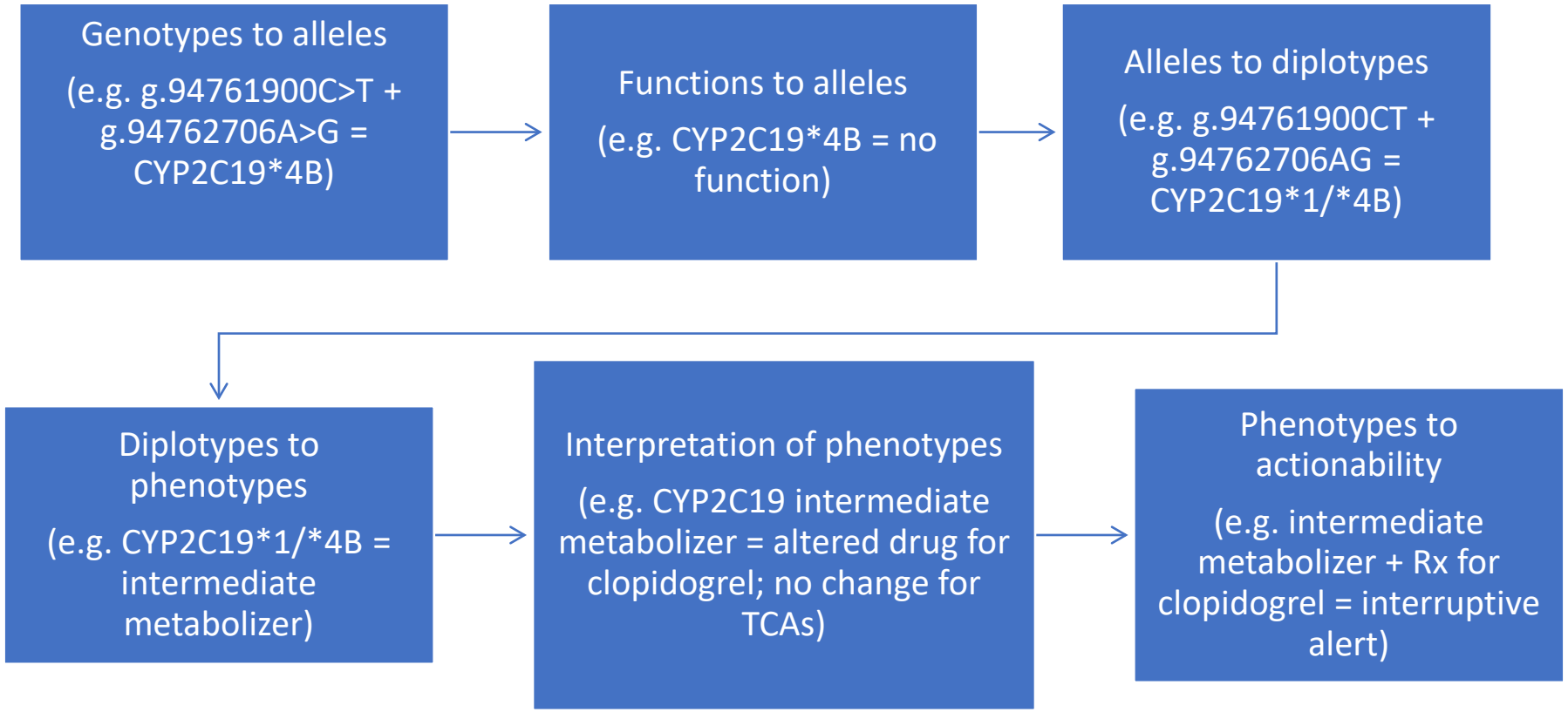
- Commentary: TE Klein, MD Ritchie. PharmCAT: A Pharmacogenomics Clinical Annotation Tool. Clinical Pharmacology & Therapeutics (2018) 104(1):19-22.
- Methods paper: K Sangkuhl & M Whirl-Carrillo, et al. Pharmacogenomics Clinical Annotation Tool (PharmCAT). Clinical Pharmacology & Therapeutics (2020) 107(1):203-210.

Motivation for PharmCAT

To automate the annotation of .vcf files with the appropriate haplotypes or diplotypes from the CPIC guideline genes*, and generate a report with the corresponding CPIC guideline prescribing recommendations

* Future versions of PharmCAT may include PGx information as well

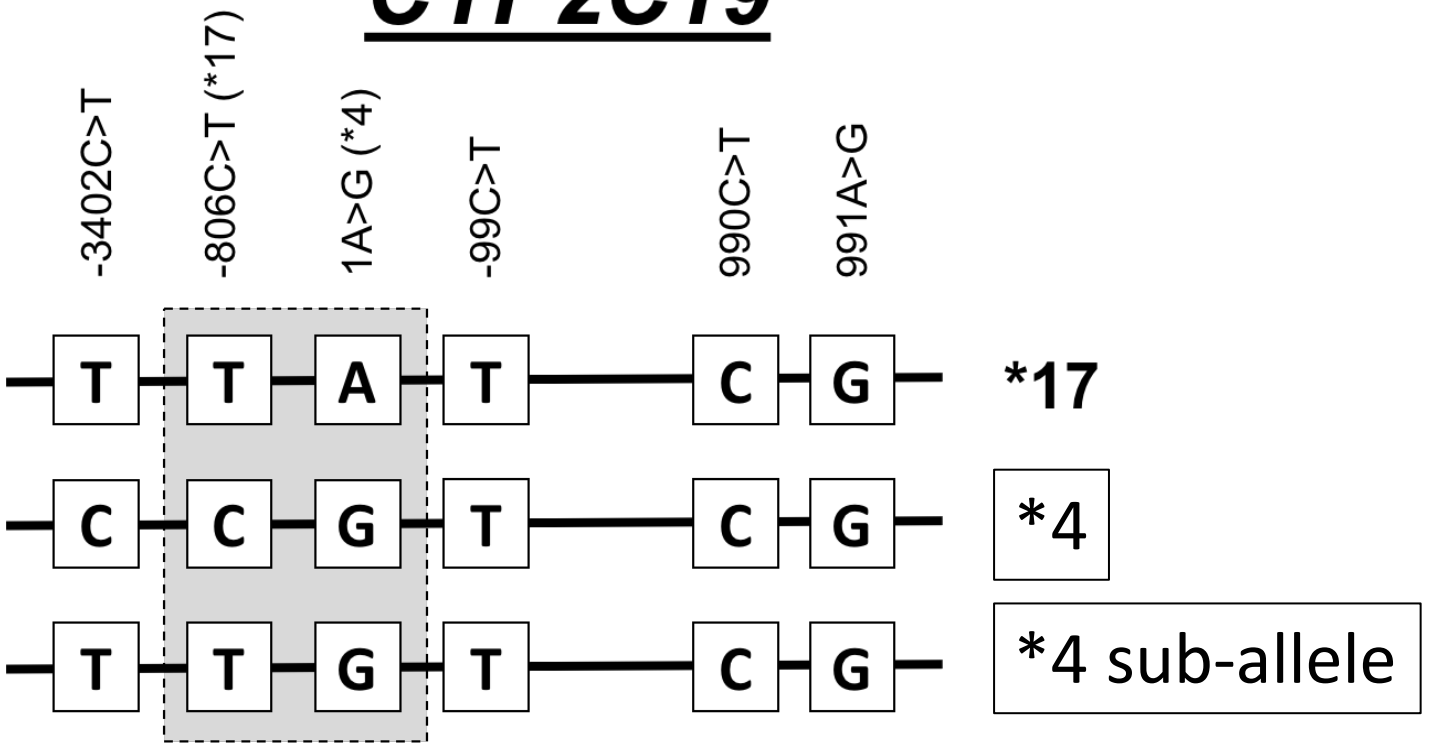
CPIC Tables:
Genetic Test Results -> Actionability



https://cpicpgx.org/guidelines/
https://www.pharmgkb.org/page/cyp2c19RefMaterials

Why is This Not Simple?

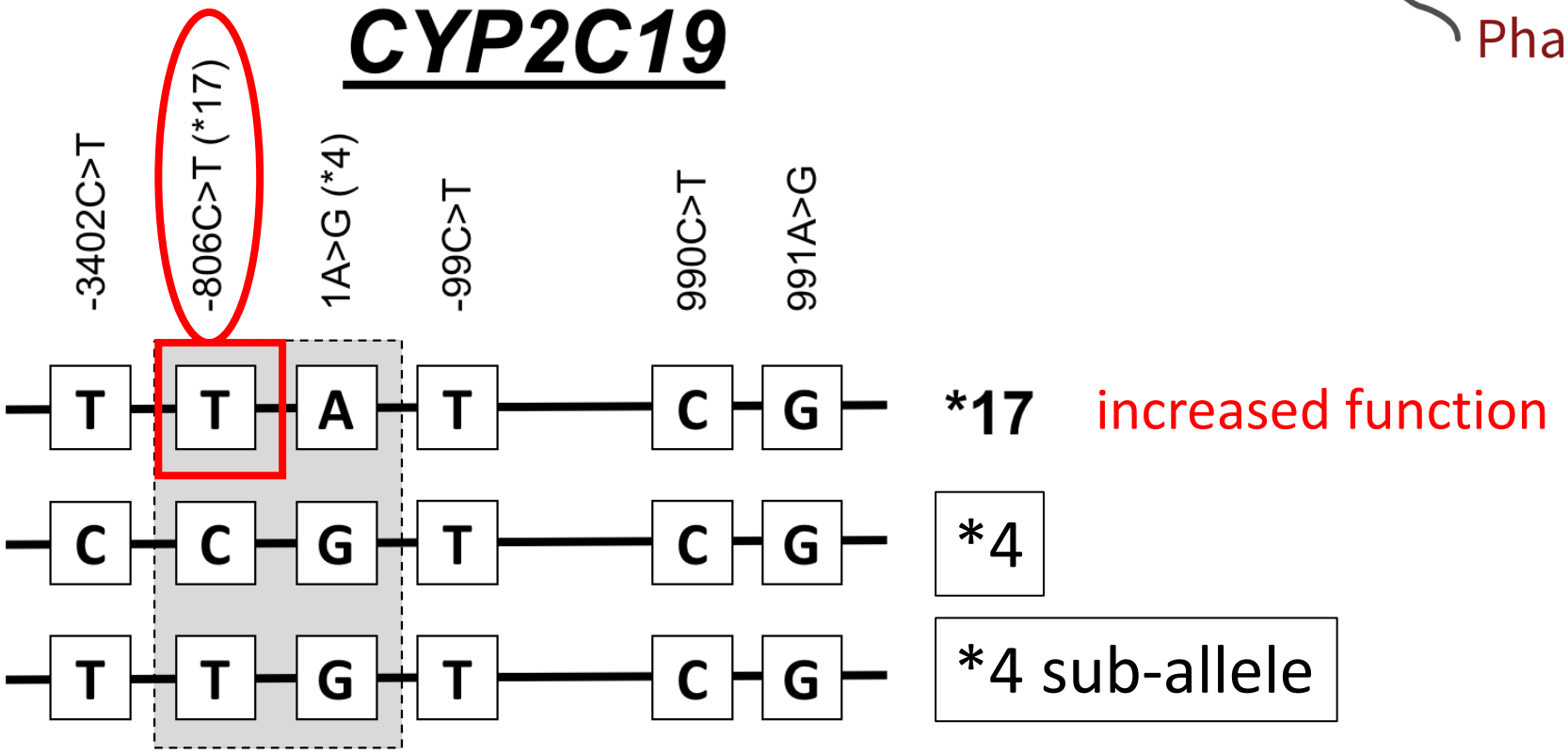
- Good quality sequence, phased – multiple options for diplotype calling
  - Many (most) don't map to phenotype or action
- Low quality sequence, in a few cases exomes, and snp chip, unphased
  - Haplotype assignment and diplotype assignment can be tricky
  - Assigned diplotypes can affect **phenotype** and **recommendations**
  - GeT-RM/CDC project sent same samples to different labs and got different results because of what was tested for – PharmCAT wants to minimize this
    - No caller can change what is tested for



Three different haplotype assignments for *CYP2C19* are shown

*17, *4, and a *4 sub-allele

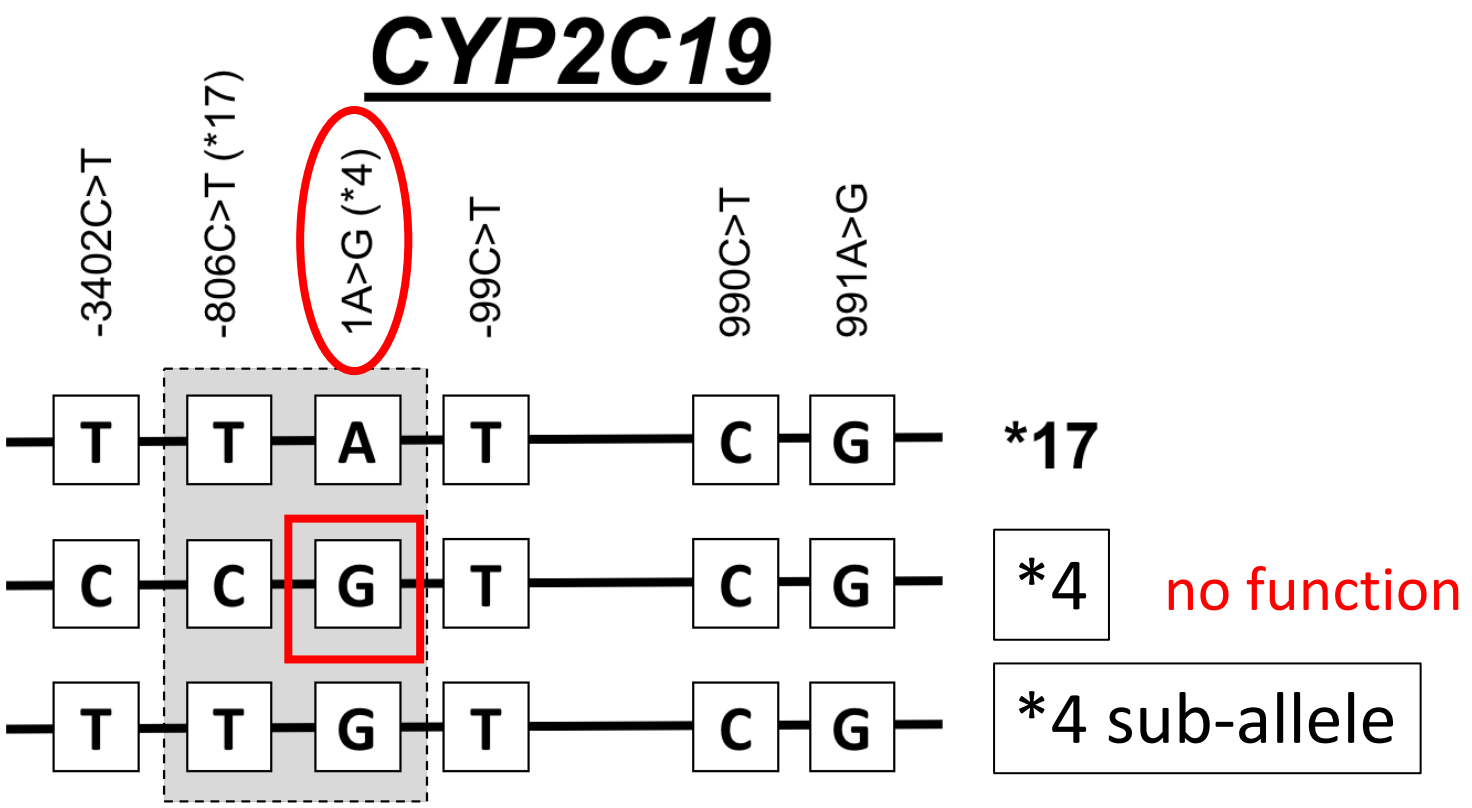
These are defined based on two specific positions (-806C>T and 1A>G)



Three different haplotype assignments for *CYP2C19* are shown

*17, *4, and a *4 sub-allele

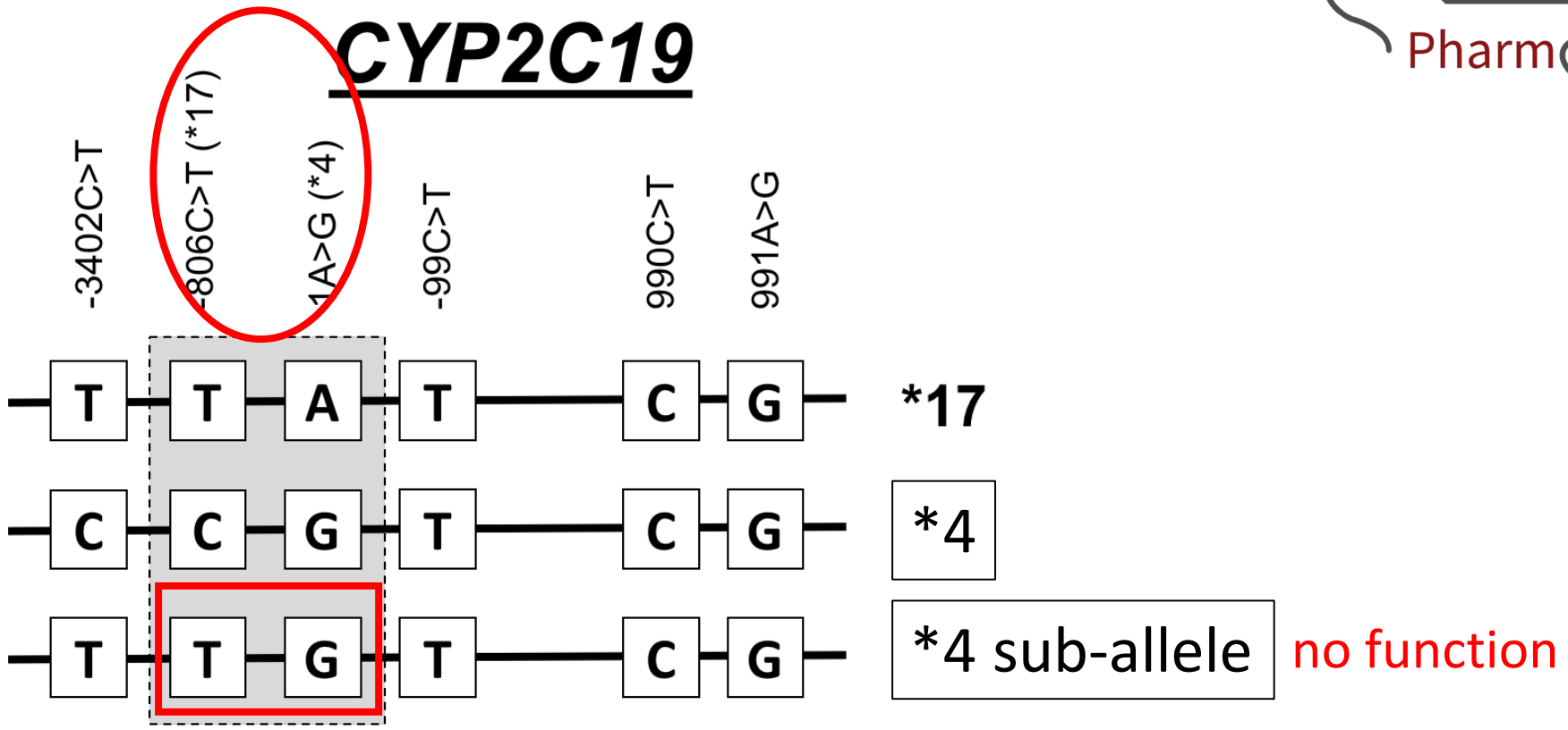
These are defined based on two specific positions (-806C>T and 1A>G)



Three different haplotype assignments for *CYP2C19* are shown

*17, *4, and a *4 sub-allele

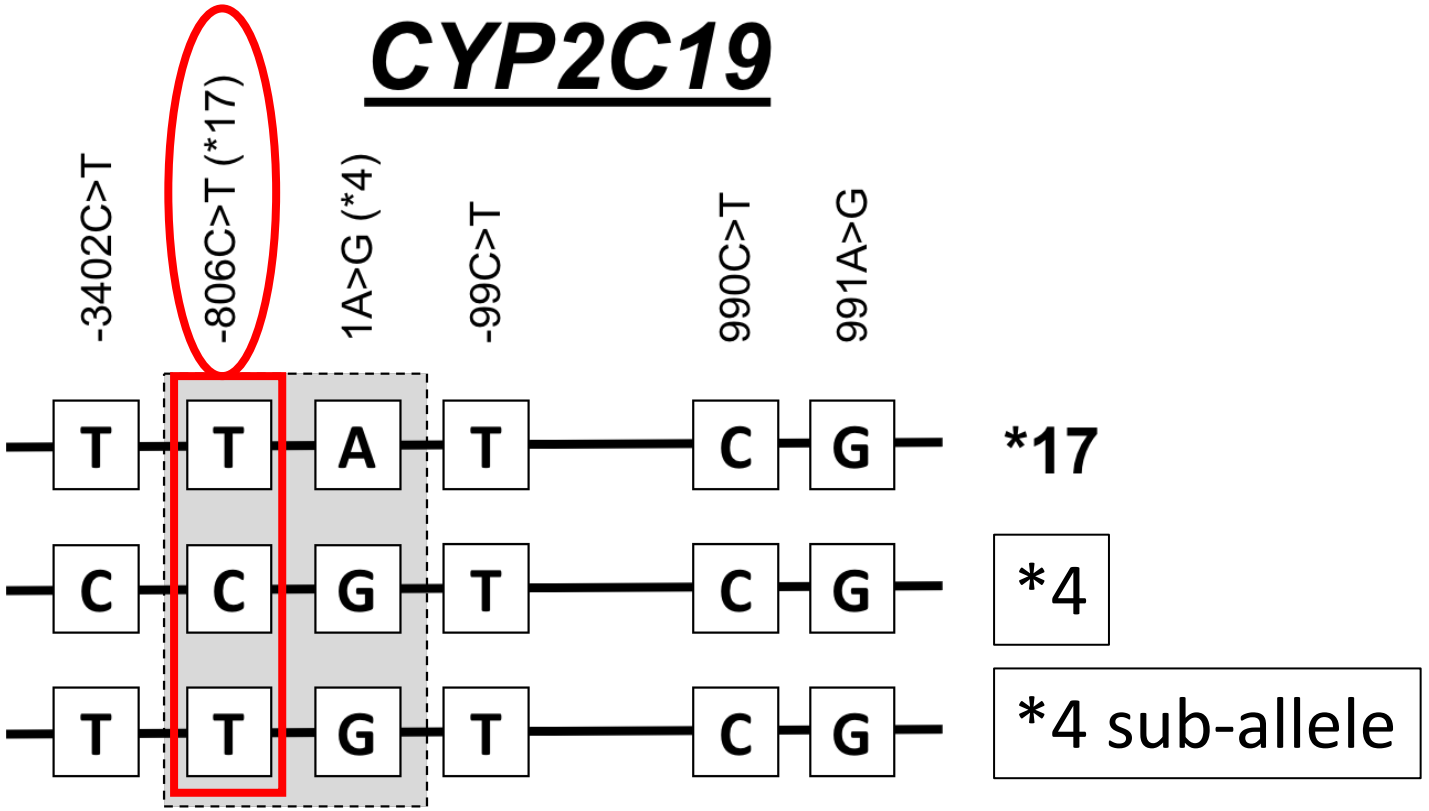
These are defined based on two specific positions (-806C>T and 1A>G)



Three different haplotype assignments for *CYP2C19* are shown

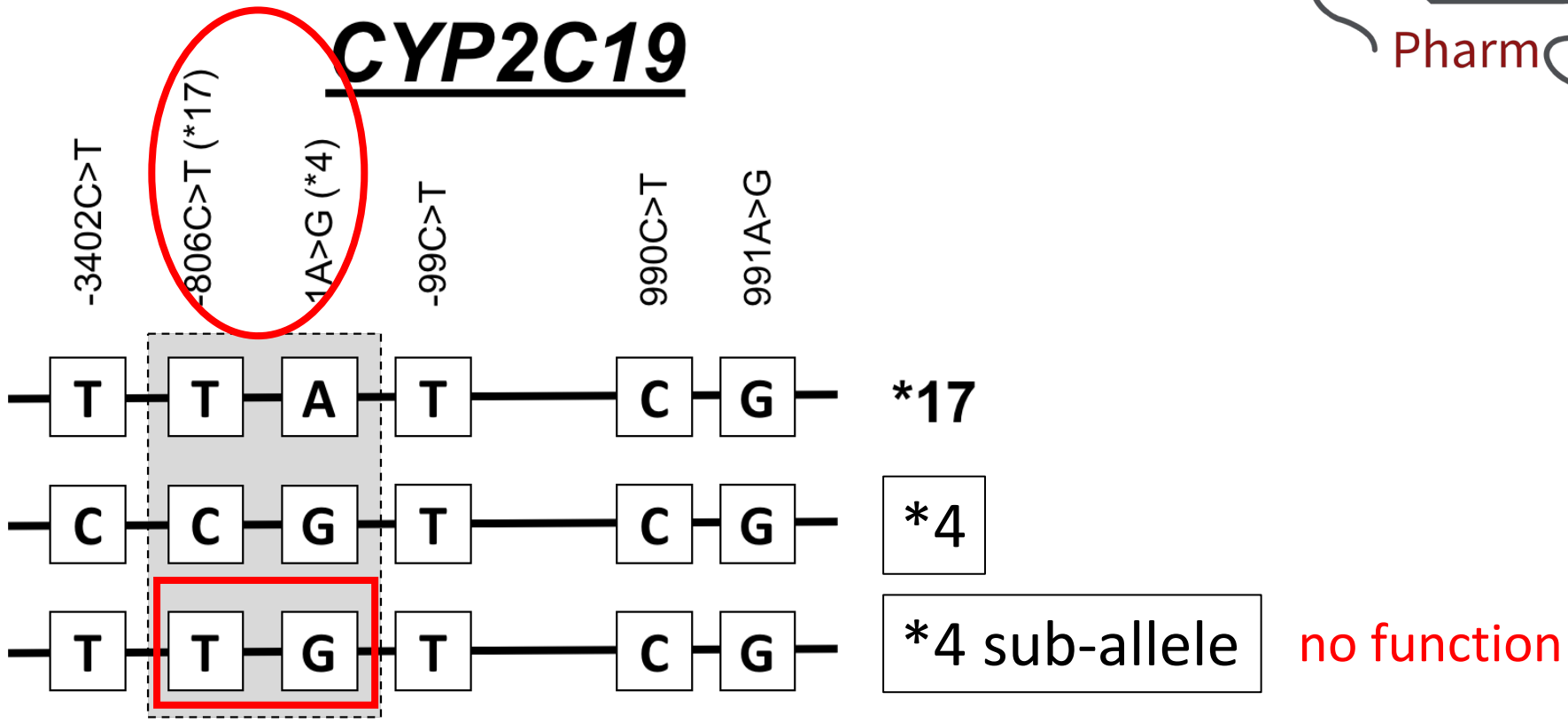
*17, *4, and a *4 sub-allele

These are defined based on two specific positions (-806C>T and 1A>G)



If both variants are present, but only the base at position -806 is reported:

Assigned diplotype = *1/*17 = RM phenotype



If both variants are present, and reported:

Are these variants in cis? -> Assigned diplotype = *1/*4 (sub-allele)

Are these variants in trans? -> Assigned diplotype = *4/*17

Both possible diplotypes are mapped to IM phenotype

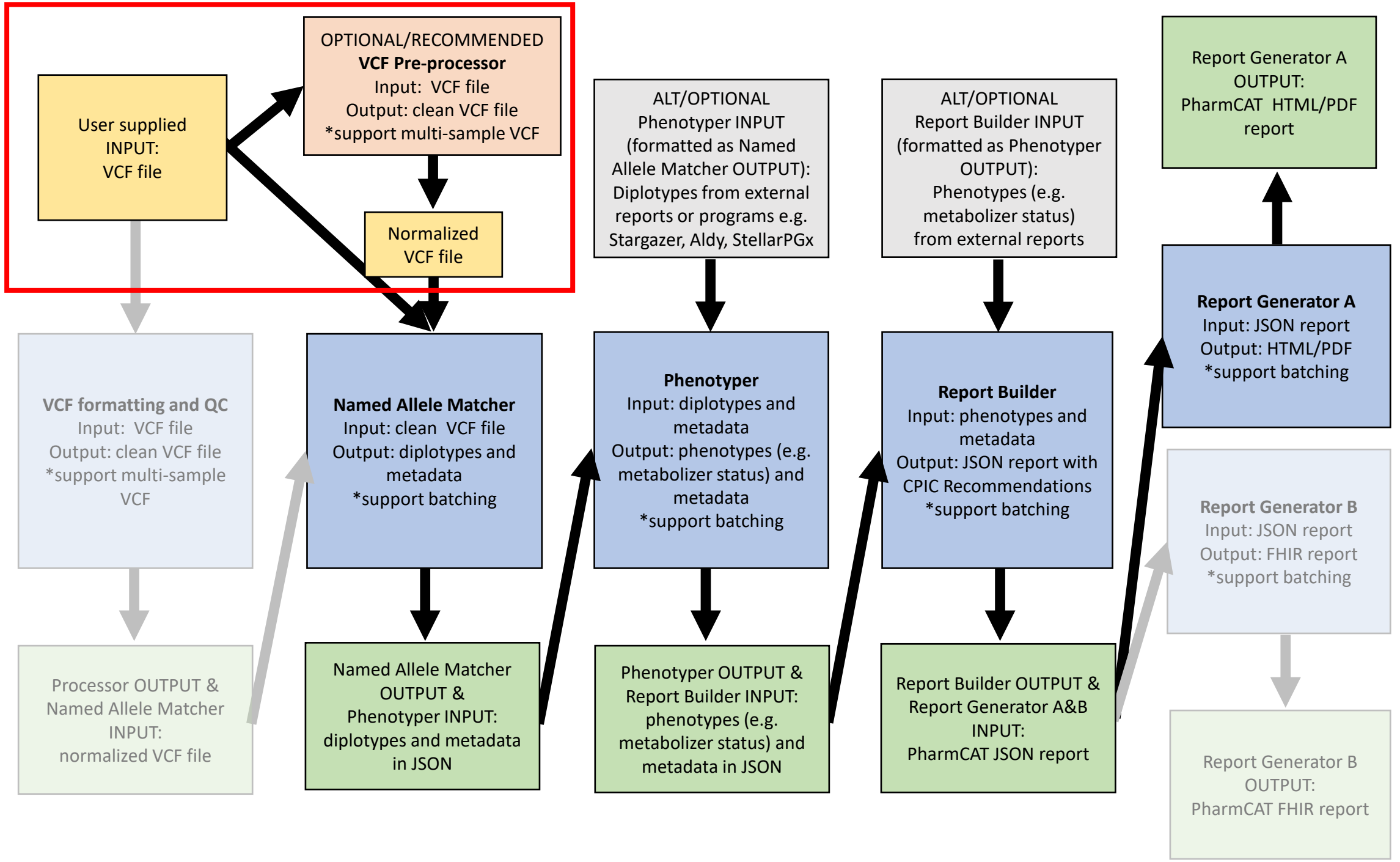
(as opposed to the RM phenotype for *1/*17)

Why Does This Matter?

- Haplotype and diplotype calling can affect **phenotype** and **recommendations**
- No tool can change what is tested for
- Need for awareness and transparency in reporting about what is tested and what is found
  - If the genotype information is there - use it
  - If the genotype information is missing – say so, be very clear what cannot be called
  - If there is >1 possibility - say so
  - If the result is defaulting to *1 because no variants are found – say what you tested for, importantly that person may NOT be *1

Pharmacogenomics Clinical Annotation Tool (PharmCAT)

- End-to-end
  - automate the annotation of .vcf files with the appropriate genotypes or diplotypes and generate a report, starting with CPIC guideline prescribing recommendations
- Research purposes may not need recommendations
- Some already have diplotypes and/or phenotypes
  - Alternate algorithm
  - Genetic test report
- Modularize tool for multiple uses


User supplied INPUT: VCF file
OPTIONAL/RECOMMENDED
VCF Pre-processor
Input: VCF file
Output: clean VCF file
*support multi-sample VCF
Normalized VCF file
ALT/OPTIONAL Phenotyper INPUT (formatted as Named Allele Matcher OUTPUT): Diplotypes from external reports or programs e.g. Stargazer, Aldy, StellarPGx
ALT/OPTIONAL Report Builder INPUT (formatted as Phenotyper OUTPUT): Phenotypes (e.g. metabolizer status) from external reports
Report Generator A OUTPUT: PharmCAT HTML/PDF report
VCF formatting and QC
Input: VCF file
Output: clean VCF file
*support multi-sample VCF
Named Allele Matcher
Input: clean VCF file
Output: diplotypes and metadata
*support batching
Phenotyper
Input: diplotypes and metadata
Output: phenotypes (e.g. metabolizer status) and metadata
*support batching
Report Builder
Input: phenotypes and metadata
Output: JSON report with CPIC Recommendations
*support batching
Report Generator A
Input: JSON report
Output: HTML/PDF
*support batching
Report Generator B
Input: JSON report
Output: FHIR report
*support batching
Processor OUTPUT & Named Allele Matcher INPUT: normalized VCF file
Named Allele Matcher OUTPUT & Phenotyper INPUT: diplotypes and metadata in JSON
Phenotyper OUTPUT & Report Builder INPUT: phenotypes (e.g. metabolizer status) and metadata in JSON
Report Builder OUTPUT & Report Generator A&B INPUT: PharmCAT JSON report
Report Generator B OUTPUT: PharmCAT FHIR report

VCF File Requirements

- Build version GRCh38
- Allele calls at PharmCAT positions must be provided including reference alleles; missing positions are assumed unknown to the user and are not used by PharmCAT
- Quality of vcf file is the responsibility of the user
- Variants should be in expected formats
- 1 VCF per sample

```
##fileformat=VCFv4.1
##fileDate=2015-08-04
##source=IlluminaPlatinumGenomes, version: hg38_2.0.1
##reference=hg38
##FORMAT=<ID=GT,Number=1,Type=String,Description="Genotype">
##FILTER=<ID=PASS,Description="All filters passed">
#CHROM  POS  ID REF  ALT  QUAL FILTER INFO FORMAT NA12878
chr10 94760676  . C . . PASS  assume-default  GT  0/0
chr10 94760686  rs111490789 C . . PASS  assume-default  GT  0/0
chr10 94761267  rs17878739  T . . PASS  assume-default  GT  0/0
chr10 94761665  rs7902257 G . . PASS  assume-default  GT  0/0
chr10 94761900  rs12248560  C T . PASS  s17 GT  0/1
chr10 94762693  rs367543001 G . . PASS  assume-default  GT  0/0
chr10 94762706  rs28399504  A . . PASS  assume-default  GT  0/0
chr10 94762712  rs367543002 C . . PASS  assume-default  GT  0/0
chr10 94762715  rs367543003 T . . PASS  assume-default  GT  0/0
chr10 94762755  rs55752064  T . . PASS  assume-default  GT  0/0
chr10 94762760  rs17882687  A . . PASS  assume-default  GT  0/0
chr10 94762788  . A . . PASS  assume-default  GT  0/0
chr10 94762856  . A . . PASS  assume-default  GT  0/0
chr10 94775106  rs145328984 C . . PASS  assume-default  GT  0/0
```

PharmCAT Provides a VCF Preprocessing Tool to Take Care of Real-World VCF Format Issues

Pharmcat

Input VCF

```
# A proper bgzip compressed VCF file ending with a suffix '.vcf.gz'
#CHROM  POS     ID      REF     ALT     QUAL    FILTER  INFO    FORMAT  SAMPLE_1
2       233760233       .       C       CAT     .       PASS    .       GT      1/0
2       233760233       .       CAT     C       .       PASS    .       GT      0/0
2       233760233       .       C       CATAT   .       PASS    .       GT      0/1
7       117548628       .       GTTTTTTA        GTTTTTA .       PASS    .       GT      0/1
7       117548628       .       GTT     G       .       PASS    .       GT      0/1
```

PharmCAT
VCF Preprocessing module

- ✓ VCF normalization (left-aligned, parsimonious format)
- ✓ Conversion to PharmCAT expected multi-allelic format
- ✓ Output PharmCAT-ready VCFs separately for each sample
- ✓ Automatically download the GRCh38/hg38 fasta and index files from the NIH FTP site if not provided
- ✓ Option to process a selected subset of samples

(Pending) Provide an option to convert ./. to 0/0

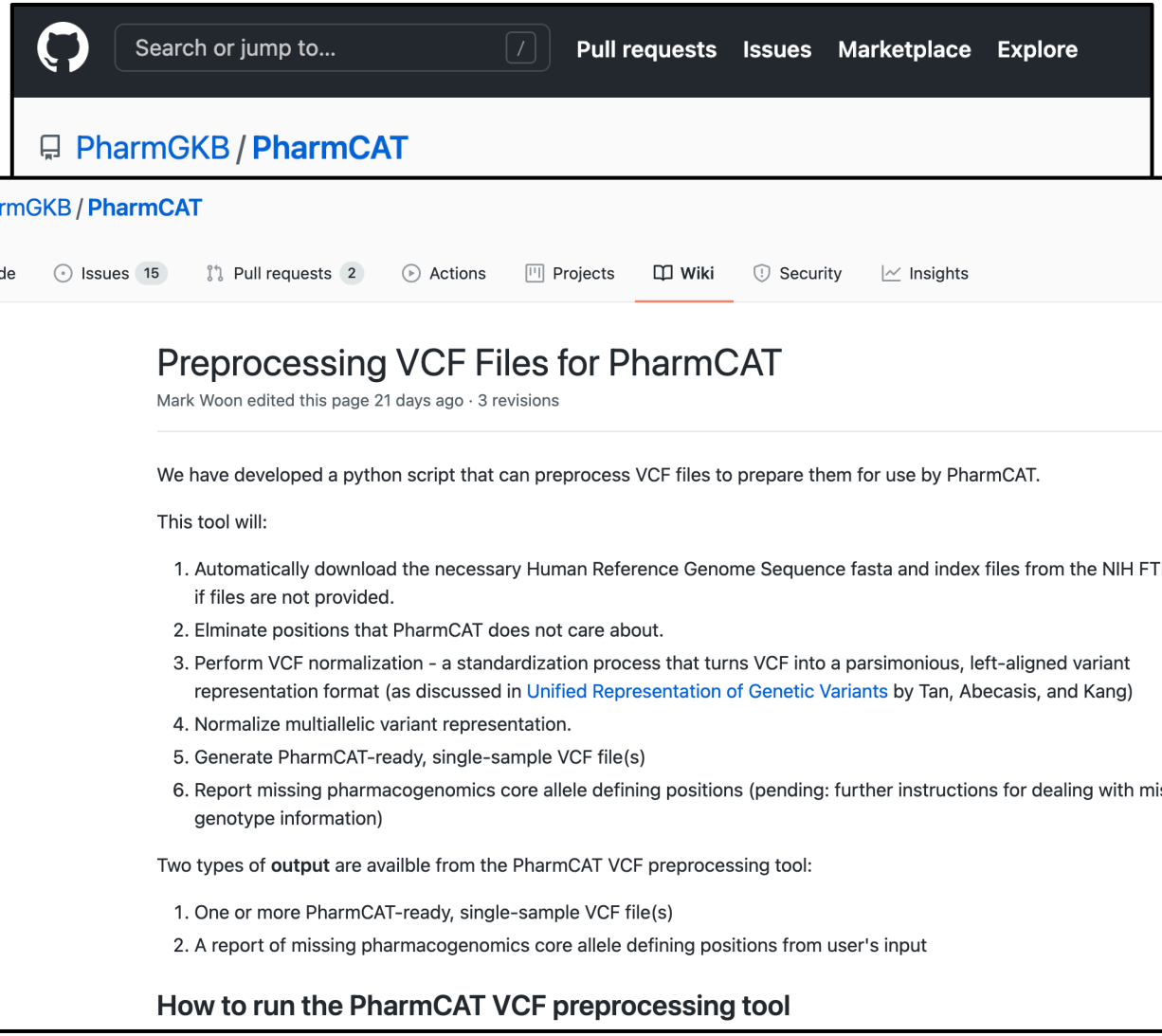
PharmCAT-ready VCF(s) for each sample

```
# A PharmCAT-ready VCF output from the PharmCAT VCF preprocessing tool
#CHROM  POS     ID      REF     ALT     QUAL    FILTER  INFO    FORMAT  PharmCAT
chr2    233760233       .       CAT     CATAT,C,CATATAT .       PASS    AC=1,0,1;AN=2   GT      1/3
chr7    117548628       .       GTT     G       .       PASS    AC=1;AN=2       GT      0/1
```

Report of missing positions

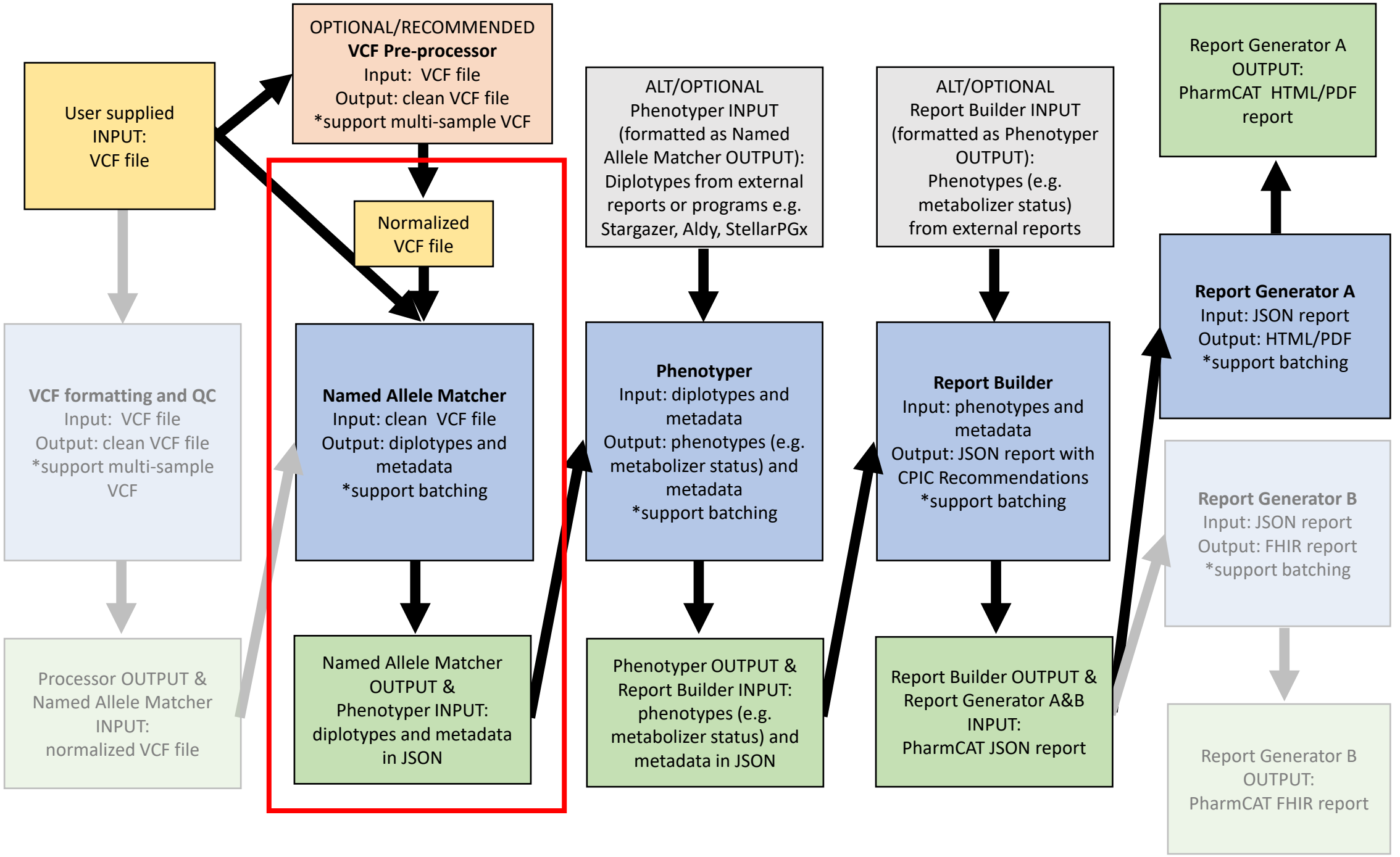
```
# A report of missing PGx allele defining positions from user's input
#CHROM  POS     ID      REF     ALT     QUAL    FILTER  INFO    FORMAT  PharmCAT
chr1    97078987        rs114096998     G       .       .       PASS    PX      GT      0/0
chr1    97078993        rs148799944     C       .       .       PASS    PX      GT      0/0
chr1    97079005        rs140114515     C       .       .       PASS    PX      GT      0/0
chr1    97079071        rs1801268       C       A       .       PASS    PX=DPYD:Reference[10]isC,DPYD:c.2983G>T[1]isA   GT      0/0
chr1    97079076        rs139459586     A       .       .       PASS    PX      GT      0/0
<truncated>
```

PharmCAT VCF Preprocessing Tool Is Publicly Available on GitHub



https://github.com/PharmGKB/PharmCAT

✓ Instructions and tutorial of the preprocessing tool are available on the GitHub wiki page.


User supplied INPUT: VCF file
OPTIONAL/RECOMMENDED
VCF Pre-processor
Input: VCF file
Output: clean VCF file
*support multi-sample VCF
Normalized VCF file
ALT/OPTIONAL Phenotyper INPUT (formatted as Named Allele Matcher OUTPUT): Diplotypes from external reports or programs e.g. Stargazer, Aldy, StellarPGx
ALT/OPTIONAL Report Builder INPUT (formatted as Phenotyper OUTPUT): Phenotypes (e.g. metabolizer status) from external reports
Report Generator A OUTPUT: PharmCAT HTML/PDF report
VCF formatting and QC
Input: VCF file
Output: clean VCF file
*support multi-sample VCF
Named Allele Matcher
Input: clean VCF file
Output: diplotypes and metadata
*support batching
Phenotyper
Input: diplotypes and metadata
Output: phenotypes (e.g. metabolizer status) and metadata
*support batching
Report Builder
Input: phenotypes and metadata
Output: JSON report with CPIC Recommendations
*support batching
Report Generator A
Input: JSON report
Output: HTML/PDF
*support batching
Report Generator B
Input: JSON report
Output: FHIR report
*support batching
Processor OUTPUT & Named Allele Matcher INPUT: normalized VCF file
Named Allele Matcher OUTPUT & Phenotyper INPUT: diplotypes and metadata in JSON
Phenotyper OUTPUT & Report Builder INPUT: phenotypes (e.g. metabolizer status) and metadata in JSON
Report Builder OUTPUT & Report Generator A&B INPUT: PharmCAT JSON report
Report Generator B OUTPUT: PharmCAT FHIR report

Step 1

```
1  ##fileformat=VCFv4.1
2  ##fileDate=2015-08-04
3  ##source=IlluminaPlatinumGenomes
4  ##reference=hg38
5  ##FORMAT=<ID=GT,Number=1,Type=Str
6  ##FILTER=<ID=PASS,Description="A
7  #CHROM  POS  ID REF  ALT  QUAL FI
8  chr10 94760676  . C . . PASS  ass
9  chr10 94760686  rs111490789 C . .
10 chr10 94761267  rs17878739  T . .
11 chr10 94761665  rs7902257 G . . P
12 chr10 94761900  rs12248560  C T .
13 chr10 94762693  rs367543001 G . .
```

User supplies a VCF file

+

GENE: CYP2C19				
Nucleotide change to gene from http://www.pharmvar.org	-806C>T	1A>G	7C>T	
Effect on protein (NP_000760.1)	5' region	M1V	P3S	
Position at NC_000010.11 (Homo sapiens chromosome 10, GRCh38.p2)	g.94761900C>T	g.94762706A>G	g.94762712C>T	g.
Position at NG_008384.3 (CYP2C19 RefSeqGene; forward relative to chromosome)	g.4220C>T	g.5026A>G	g.5032C>T	
rsID	rs12248560	rs28399504	rs367543002	r
CYP2C19 Allele				
*38	C	A	C	
*1				
*2				
*3				
*4	Y	G		
*5				
*6				
*7				
*8				
*9				
*10				
*11				
*12				
*13				

PharmCAT provides allele definition files Based on CPIC/PharmGKB and PharmVar core alleles

=

Gene Diplotypes And Variant Genotypes For CPIC Genes and Variants

Named Allele Matcher

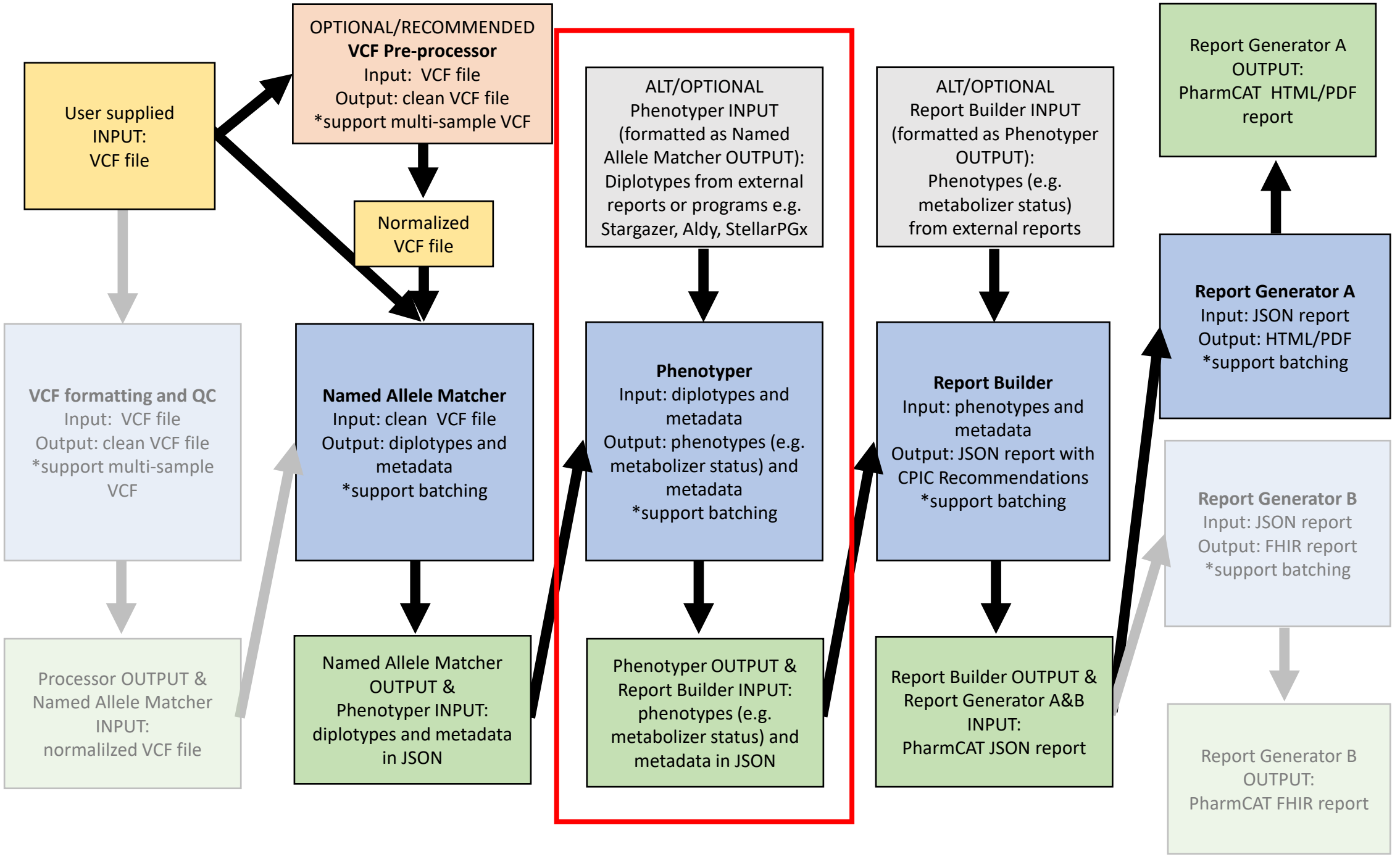
CYP2C19

- *1/*2 (40)

Definition Position	94760676	94760686	94761267	94761665	94761900	94762693	94762706	94762712	94762715	94762755	94762760	94762788	94762856	94775
	rs113164681	rs111490789	rs17878739	rs7902257	rs12248560	rs367543001	rs28399504	rs367543002	rs367543003	rs55752064	rs17882687			rs1453
VCF Position	94760676	94760686	94761267	94761665	94761900	94762693	94762706	94762712	94762715	94762755	94762760	94762788	94762856	94775
VCF REF,ALTs	C	C	T	G	C	G	A	C	T	T	A	A	A	C
VCF Call	C|C	C|C	T|T	G|G	C|C	G|G	A|A	C|C	T|T	T|T	A|A	A|A	A|A	C|
*1	C	C	T	G	C	G	A	C	T	T	A	A	A	C
	C	C	T	G	C	G	A	C	T	T	A	A	A	C
*2	C	C	T	G	C	G	A	C	T	T	A	A	A	C
	C	C	T	G	C	G	A	C	T	T	A	A	A	C

Algorithms to match allele definitions to variant call data from vcf input:

- If data is unphased, generate all possible combinations of genotypes for the positions of interest.
- Attempt to match each combination to a named allele.
- If there are matches, and data is unphased, try to build diplotypes by making sure that the genotype combinations are possible – take longest match.


User supplied INPUT: VCF file
OPTIONAL/RECOMMENDED
VCF Pre-processor
Input: VCF file
Output: clean VCF file
*support multi-sample VCF
Normalized VCF file
ALT/OPTIONAL Phenotyper INPUT (formatted as Named Allele Matcher OUTPUT): Diplotypes from external reports or programs e.g. Stargazer, Aldy, StellarPGx
ALT/OPTIONAL Report Builder INPUT (formatted as Phenotyper OUTPUT): Phenotypes (e.g. metabolizer status) from external reports
Report Generator A OUTPUT: PharmCAT HTML/PDF report
VCF formatting and QC
Input: VCF file
Output: clean VCF file
*support multi-sample VCF
Named Allele Matcher
Input: clean VCF file
Output: diplotypes and metadata
*support batching
Phenotyper
Input: diplotypes and metadata
Output: phenotypes (e.g. metabolizer status) and metadata
*support batching
Report Builder
Input: phenotypes and metadata
Output: JSON report with CPIC Recommendations
*support batching
Report Generator A
Input: JSON report
Output: HTML/PDF
*support batching
Processor OUTPUT & Named Allele Matcher INPUT: normalilzed VCF file
Named Allele Matcher OUTPUT & Phenotyper INPUT: diplotypes and metadata in JSON
Phenotyper OUTPUT & Report Builder INPUT: phenotypes (e.g. metabolizer status) and metadata in JSON
Report Builder OUTPUT & Report Generator A&B INPUT: PharmCAT JSON report
Report Generator B
Input: JSON report
Output: FHIR report
*support batching
Report Generator B OUTPUT: PharmCAT FHIR report

Step 2

Gene Diplotypes
And
Variant Genotypes
For CPIC Genes
and Variants

+

	A	D	F	G	H
1	GENE: CYP2C19				
2	Allele/cDNA/rsID	Allele Clinical Functional Status (Required)	PMID (Required)	Strength of Evidence (Required)	Summary of Findings
3	*1	Normal function	7487078, 32602114, 22027650	Definitve	CYP2C19*1 is assigned normal function based on been studied for more than 20 years and is the mos 32602114). CYP2C19*1 has repeatedly demonstr substrates and its association with normal funct
4	*2	No function	8195181, 22027650	Definitve	CYP2C19*2 is assigned no function based on defir studied for more than 20 years and its associa established (8195181, 22027650). CYP2C19*2 is d a truncated protein lacking a heme binding regio
5	*3	No function	7969038, 9103550, 22027650	Definitve	CYP2C19*3 is assigned no function based on defir studied for more than 20 years and its associa established (7969038, 9103550, 22027650). CYP codon resulting in a truncated protein lacking a hem substrate binding region, making it a nonfunctio
	*4	No function	9435198, 21358751	Limited	CYP2C19*4 is assigned no function based on limit and an in vitro study. Two subjects with CYP2

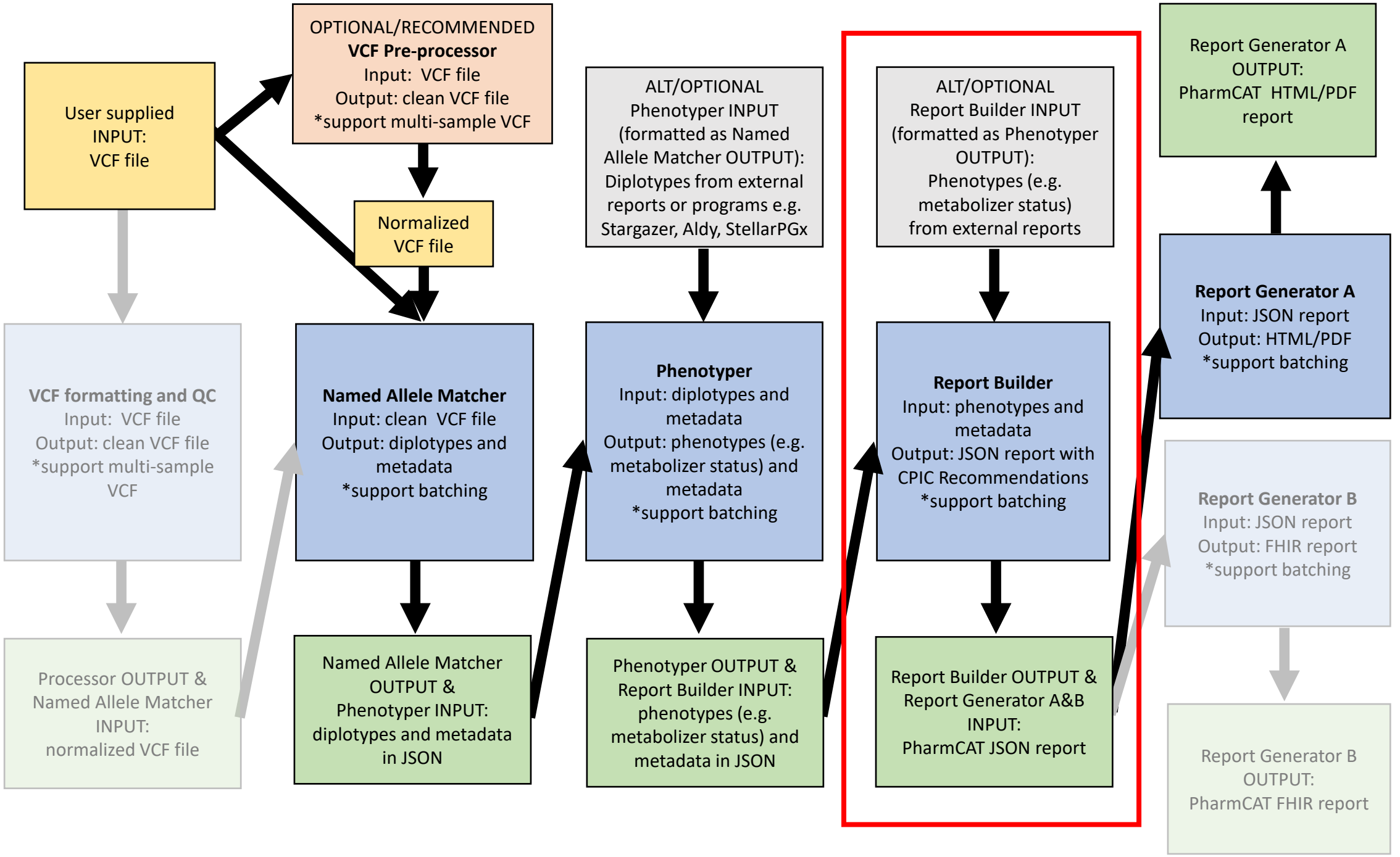
=

Phenotypes
For CPIC
Genes and
Variants

Optional:
Gene Diplotypes
From External
Program
or Genetic Report

Gene: CYP2C19			
Allele 1 Function	**Allele 2 Function**	**Phenotype**	**Description**
Increased function	Increased function	Ultrarapid Metabolizer	An individual carrying two increased
Increased function	Normal function	Rapid Metabolizer	An individual carrying one normal fu
Normal function	Normal function	Normal Metabolizer	An individual carrying two normal fu
Decreased function	Normal function	Likely Intermediate Metabolizer	An individual carrying one normal fu
Decreased function	Increased function	Likely Intermediate Metabolizer	An individual carrying one increased
Decreased function	Decreased function	Likely Intermediate Metabolizer	An individual carrying two decrease
Normal function	No function	Intermediate Metabolizer	An individual carrying one normal fu
Increased function	No function	Intermediate Metabolizer	An individual carrying one increased
Decreased function	No function	Likely Poor Metabolizer	An individual carrying one decrease
No function	No function	Poor Metabolizer	An individual carrying two no functi

PharmCAT provides allele function and
phenotype mapping based on CPIC/PharmGKB


User supplied INPUT: VCF file
OPTIONAL/RECOMMENDED
VCF Pre-processor
Input: VCF file
Output: clean VCF file
*support multi-sample VCF
Normalized VCF file
VCF formatting and QC
Input: VCF file
Output: clean VCF file
*support multi-sample VCF
Processor OUTPUT & Named Allele Matcher INPUT: normalized VCF file
Named Allele Matcher
Input: clean VCF file
Output: diplotypes and metadata
*support batching
Named Allele Matcher OUTPUT & Phenotyper INPUT: diplotypes and metadata in JSON
ALT/OPTIONAL Phenotyper INPUT (formatted as Named Allele Matcher OUTPUT): Diplotypes from external reports or programs e.g. Stargazer, Aldy, StellarPGx
Phenotyper
Input: diplotypes and metadata
Output: phenotypes (e.g. metabolizer status) and metadata
*support batching
Phenotyper OUTPUT & Report Builder INPUT: phenotypes (e.g. metabolizer status) and metadata in JSON
ALT/OPTIONAL Report Builder INPUT (formatted as Phenotyper OUTPUT): Phenotypes (e.g. metabolizer status) from external reports
Report Builder
Input: phenotypes and metadata
Output: JSON report with CPIC Recommendations
*support batching
Report Builder OUTPUT & Report Generator A&B INPUT: PharmCAT JSON report
Report Generator A OUTPUT: PharmCAT HTML/PDF report
Report Generator A
Input: JSON report
Output: HTML/PDF
*support batching
Report Generator B
Input: JSON report
Output: FHIR report
*support batching
Report Generator B OUTPUT: PharmCAT FHIR report

Step 3

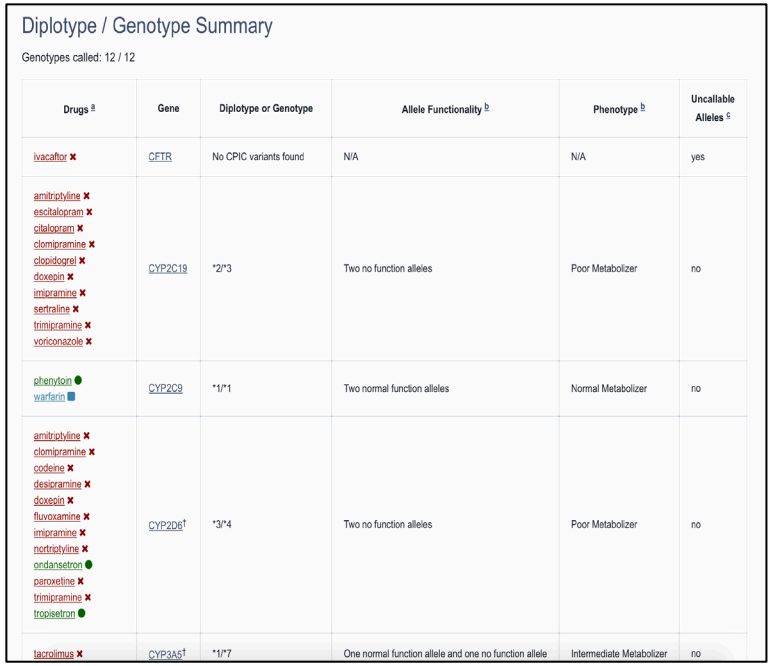
Phenotypes For CPIC Genes and Variants

+

CYP2C19 Phenotype	Implications for Phenotypic Measures	Therapeutic Recommendation	Cla Rec
CYP2C19 ultrarapid metabolizer	Decreased plasma concentrations of PPIs compared to CYP2C19 NMs; increased risk of therapeutic failure	Increase starting daily dose by 100%. Daily dose may be given in divided doses. Monitor for efficacy.	Opt
CYP2C19 rapid metabolizer	Decreased plasma concentrations of PPIs compared to CYP2C19 NMs; increased risk of therapeutic failure	Initiate standard starting daily dose. Consider increasing dose by 50-100% for the treatment of H. pylori infection and erosive esophagitis. Daily dose may be given in divided doses. Monitor for efficacy.	Mo
CYP2C19 normal metabolizer	Normal PPI metabolism; may be at increased risk of therapeutic failure compared to CYP2C19 IMs and PMs	Initiate standard starting daily dose. Consider increasing dose by 50-100% for the treatment of H. pylori infection and	Mo

PharmCAT provides CPIC guideline recommendations based on phenotype

=

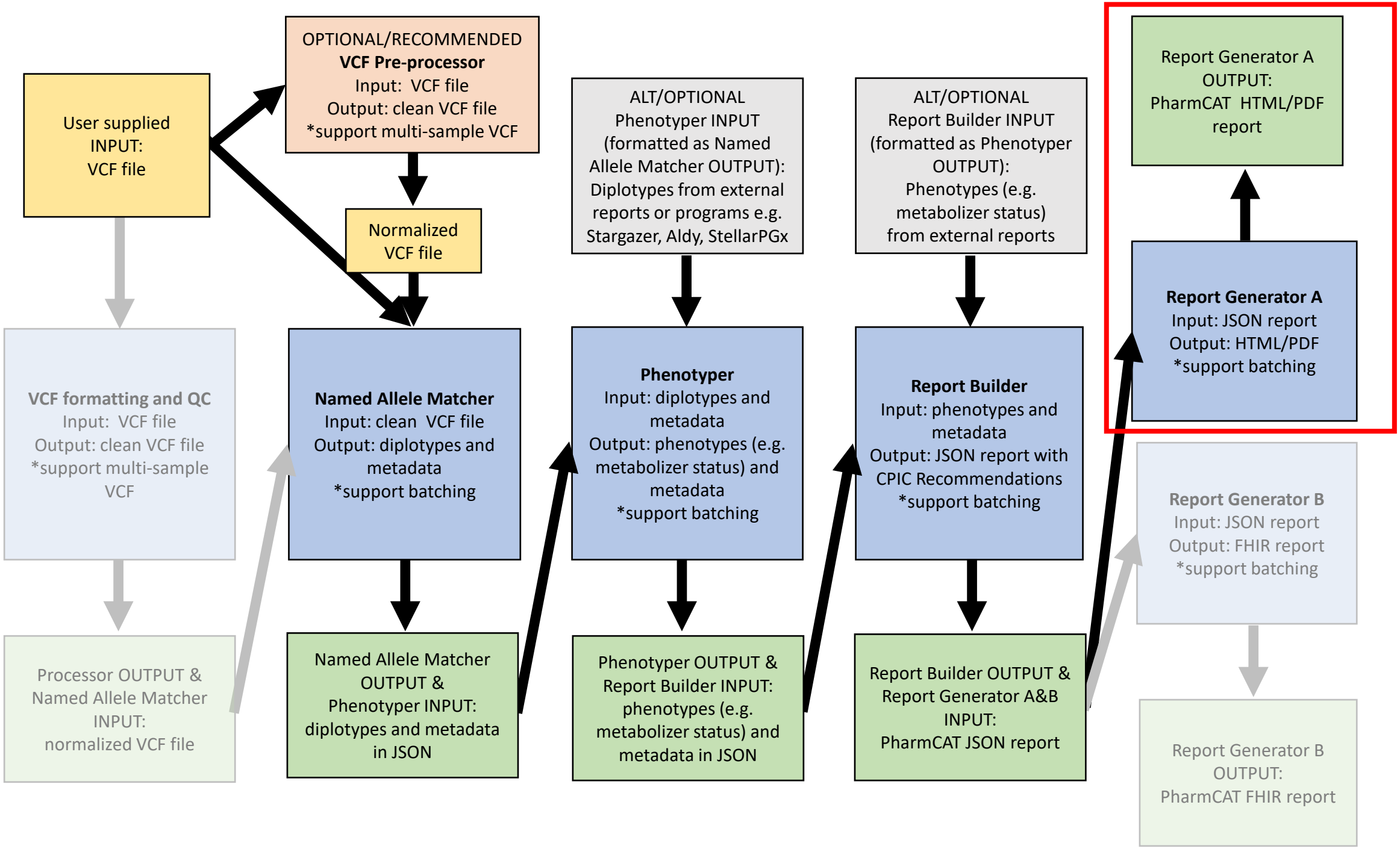


PharmCAT JSON and HTML reports

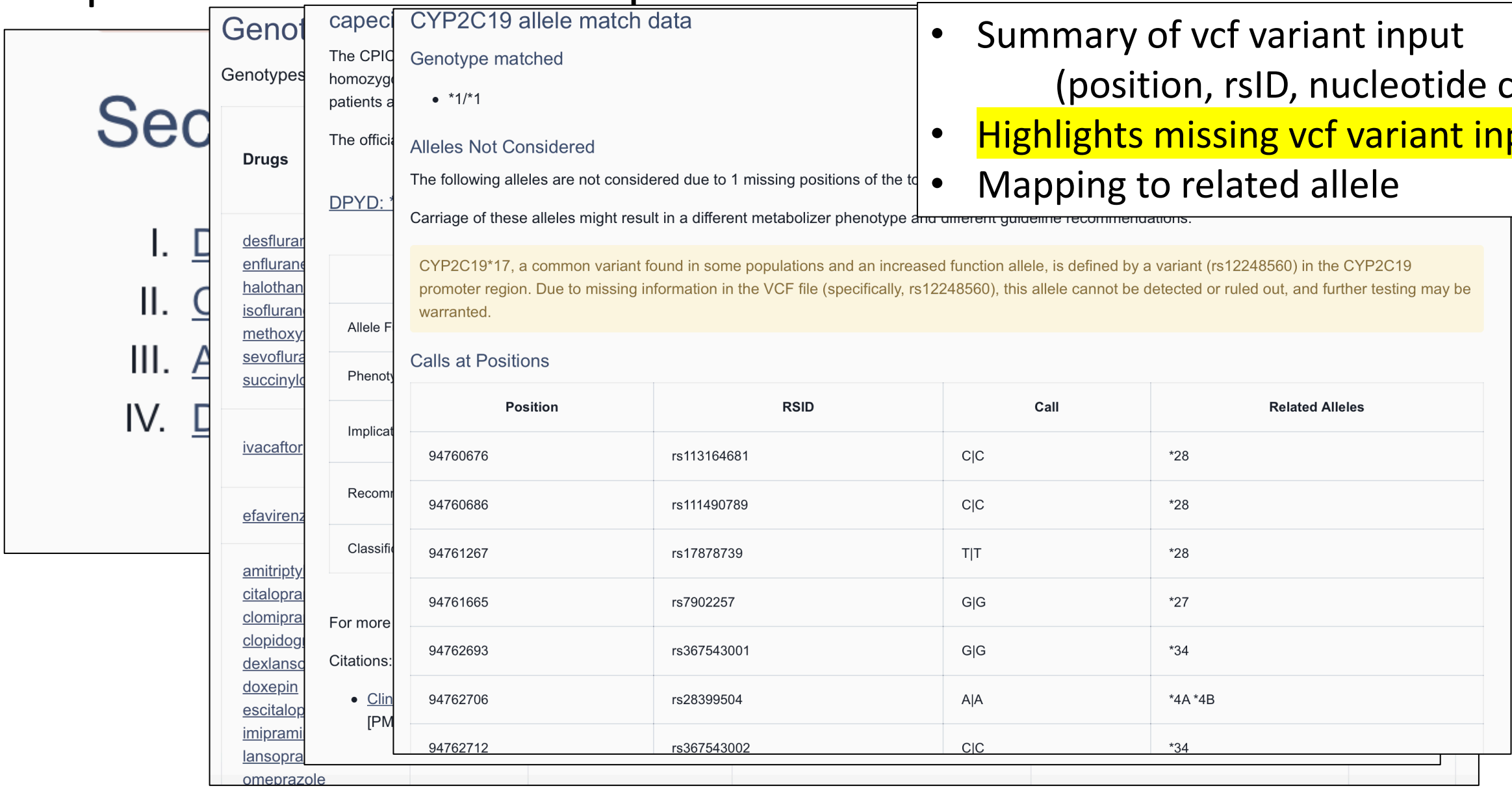
Optional: Gene Phenotypes From Genetic Report

	A	B	H	I	J	
	Name	***Match:Gene***	***Match: Guideline***	***Notes***	***Type***	***Message***
	CYP2C19 *17 check for *4B	CYP2C19			ambiguity	Warning! Haplotype matches *1
	CYP2C19 *17 missing	CYP2C19		display in case CYP2C1	gene-specific warning	CYP2C19*17, a common varian
	TPMT *1/*3A warning	TPMT	azathioprine, mercaptopurine, thioguanine		ambiguity	The assigned genotype is *1/*3
	TPMT reverse complement footnote	TPMT			footnote	The TPMT gene is on the negat
	VKORC1	VKORC1			footnote	CPIC recommendations for warf
	UGT1A1 repeat warning	UGT1A1			footnote	Some alleles in UGT1A1 are dis
	CYP3A5 *3 missing	CYP3A5	tacrolimus	display in case CYP3A5	gene-specific warning	CYP3A5*3 is a common variant
	CYP3A5 reverse complement footnote	CYP3A5			footnote	The CYP3A5 gene is on the ne
	Warfarin warning		warfarin		footnote	The CPIC warfarin guideline onl
	VKORC1 reverse complement footnote	VKORC1			footnote	The VKORC1 gene is on the ne
	DPYD reverse complement footnote	DPYD			footnote	The DPYD gene is on the negat
	IFNL3 reverse complement footnote	IFNL3			footnote	The IFNL3 gene is on the negat
	Phenytoin_HLA-B warning		phenytoin	add under the diplotype	note	The displayed recommendation
	clopidogrel table footnote		clopidogrel		footnote	Antiplatelet therapy recommend
	CYP2D6 warning	CYP2D6			note	CYP2D6 genotypes are called h

PharmCAT provides case-based messages and caveats


User supplied INPUT: VCF file
OPTIONAL/RECOMMENDED
VCF Pre-processor
Input: VCF file
Output: clean VCF file
*support multi-sample VCF
Normalized VCF file
ALT/OPTIONAL Phenotyper INPUT (formatted as Named Allele Matcher OUTPUT): Diplotypes from external reports or programs e.g. Stargazer, Aldy, StellarPGx
ALT/OPTIONAL Report Builder INPUT (formatted as Phenotyper OUTPUT): Phenotypes (e.g. metabolizer status) from external reports
Report Generator A OUTPUT: PharmCAT HTML/PDF report
VCF formatting and QC
Input: VCF file
Output: clean VCF file
*support multi-sample VCF
Named Allele Matcher
Input: clean VCF file
Output: diplotypes and metadata
*support batching
Phenotyper
Input: diplotypes and metadata
Output: phenotypes (e.g. metabolizer status) and metadata
*support batching
Report Builder
Input: phenotypes and metadata
Output: JSON report with CPIC Recommendations
*support batching
Report Generator A
Input: JSON report
Output: HTML/PDF
*support batching
Report Generator B
Input: JSON report
Output: FHIR report
*support batching
Processor OUTPUT & Named Allele Matcher INPUT: normalized VCF file
Named Allele Matcher OUTPUT & Phenotyper INPUT: diplotypes and metadata in JSON
Phenotyper OUTPUT & Report Builder INPUT: phenotypes (e.g. metabolizer status) and metadata in JSON
Report Builder OUTPUT & Report Generator A&B INPUT: PharmCAT JSON report
Report Generator B OUTPUT: PharmCAT FHIR report

Step 4: PharmCAT Report



- Summary of vcf variant input (position, rsID, nucleotide call)
- Highlights missing vcf variant input
- Mapping to related allele

Current Status: Release v1.0.0

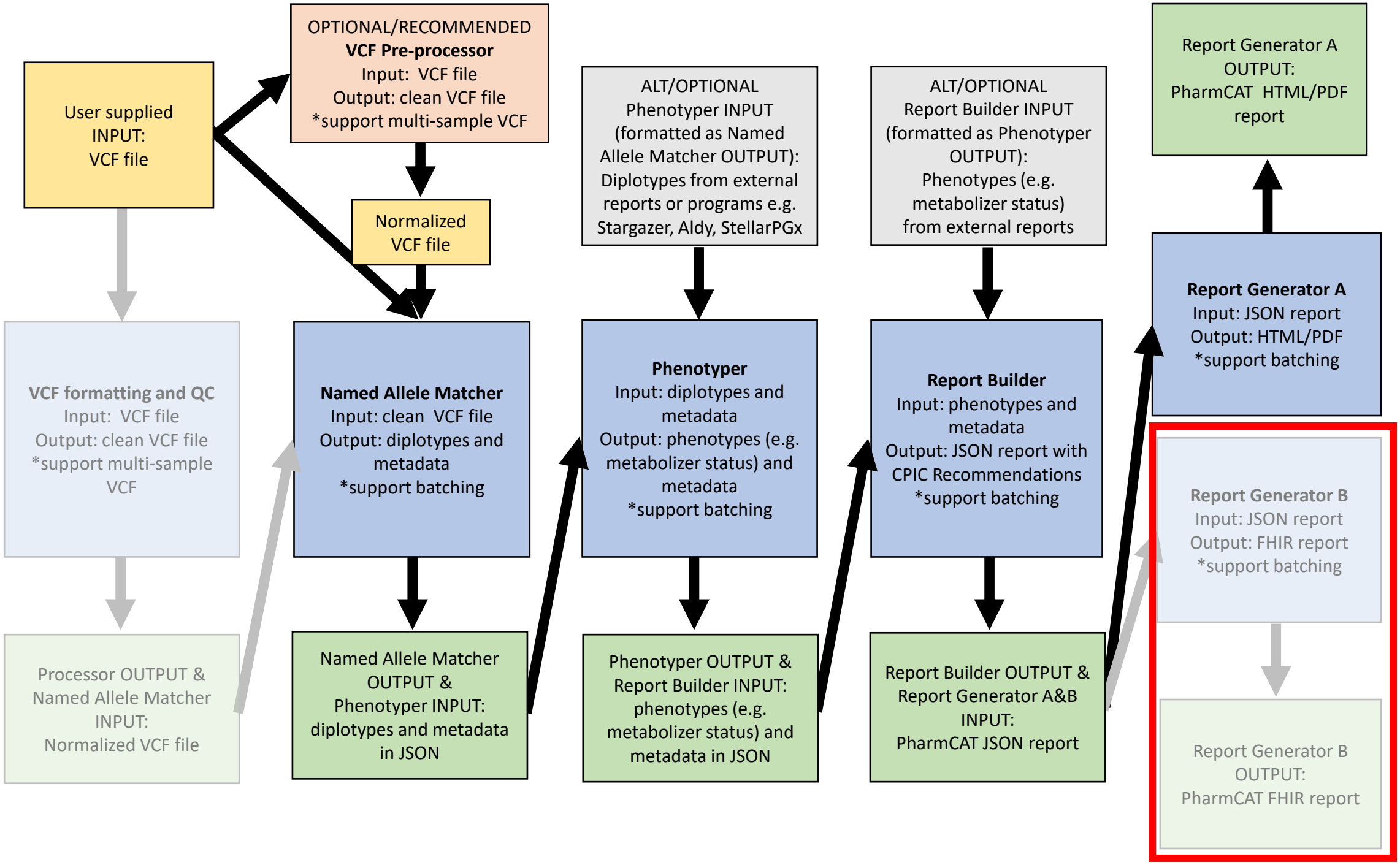
- Based on data from the CPIC DB via the API
  - Each PharmCAT release will be current with what was pulled from the CPIC DB with each release
- Supported gene and drug coverage at https://pharmcat.org/summary/
  - Note that warfarin and peg-interferon don't have recommendations in the CPIC DB
- Named Allele Matcher genes:
  - CACNA1S, CFTR, CYP2B6, CYP2C9, CYP2C19, CYP3A5, CYP4F2, DPYD, IFLN3, NUDT15, RYR1, SLCO1B1, TPMT, UGT1A1, VKORC1
- Phenotyper genes:
  - Genes above + CYP2D6, G6PD, MT-RNR1 (external source)
- Statistics based on UK Biobank samples

PharmCAT average run time: 4.3 sec/sample

Pre-processing tool average run time: 3.8 sec/sample

PharmCAT integration into EHR

- Developers of EHRs working hard to integrate genomics into clinical records
- Standards have been under development for several years
- We are leveraging these standards to integrate PharmCAT with EHRs


User supplied INPUT: VCF file
OPTIONAL/RECOMMENDED
VCF Pre-processor
Input: VCF file
Output: clean VCF file
*support multi-sample VCF
Normalized VCF file
ALT/OPTIONAL Phenotyper INPUT (formatted as Named Allele Matcher OUTPUT): Diplotypes from external reports or programs e.g. Stargazer, Aldy, StellarPGx
ALT/OPTIONAL Report Builder INPUT (formatted as Phenotyper OUTPUT): Phenotypes (e.g. metabolizer status) from external reports
Report Generator A OUTPUT: PharmCAT HTML/PDF report
VCF formatting and QC
Input: VCF file
Output: clean VCF file
*support multi-sample VCF
Named Allele Matcher
Input: clean VCF file
Output: diplotypes and metadata
*support batching
Phenotyper
Input: diplotypes and metadata
Output: phenotypes (e.g. metabolizer status) and metadata
*support batching
Report Builder
Input: phenotypes and metadata
Output: JSON report with CPIC Recommendations
*support batching
Report Generator A
Input: JSON report
Output: HTML/PDF
*support batching
Report Generator B
Input: JSON report
Output: FHIR report
*support batching
Processor OUTPUT & Named Allele Matcher INPUT: Normalized VCF file
Named Allele Matcher OUTPUT & Phenotyper INPUT: diplotypes and metadata in JSON
Phenotyper OUTPUT & Report Builder INPUT: phenotypes (e.g. metabolizer status) and metadata in JSON
Report Builder OUTPUT & Report Generator A&B INPUT: PharmCAT JSON report
Report Generator B OUTPUT: PharmCAT FHIR report

PharmCAT integration into EHR

- Developers of EHRs working hard to integrate genomics into clinical records
- Standards have been under development for several years
- We are leveraging these standards to integrate PharmCAT with EHRs
- Test integration of various modules of PharmCAT within Epic @ UPenn
- Test PharmCAT on biobank samples and return in Epic development mode @ UPenn
- Build code to integrate with Epic using HL-7 standards (with and without genomics module)
- Disseminate best practices to other institutions (e.g. Stanford)

Partial and Combination Allele Calls

- When looking at biobank/very large datasets, we have seen variant combinations not catalogued from submissions in PharmVar
  - Not novel variants, but variants in existing allele definitions found in novel combinations
- Currently, PharmCAT gives a “no call” in these situations
  - Appropriate for clinical implementation purposes
- We would like to capture what variants are seen together in these cases for phased data
  - Good for research, discovery and phenotype prediction
- Future:
  - batch processing, integrated pre-processing step
  - adding other genes/drugs/prescribing information (e.g. FDA label information)
  - identifying and predicting function of novel variants (not in this grant cycle)

Summary

- PharmCAT was created to fill a gap in the field of implementation of pharmacogenomics into clinical care
- Our philosophy is to create a resource with community input and engagement that can create more transparency in PGx reporting of results
- Future versions will include greater functionality

Acknowledgements

Stanford University

- Teri Klein
- Binglan Li
- Katrin Sangkuhl
- Ryan Whaley
- Michelle Whirl-Carrillo
- Mark Woon

University of Pennsylvania

- Marylyn Ritchie
- Yuki Bradford
- Scott Dudek
- Anastasia Lucas
- Sony Tuteja
- Anurag Verma
- Shefali Setia Verma

SAB Members

- Burns Blaxall
- Rhonda Cooper-DeHoff
- Phil Empey (Chair)
- Andrea Gaedigk
- Janina Jeff
- James Hoffman
- Houda Hachad
- Jonathan Haines
- Stuart Scott
- Casey Overby Taylor
- Sara Van Driest
- Marc Williams

NIH NHGRI

- Rongling Li

http://pharmcat.org

NHGRI U24HG010862