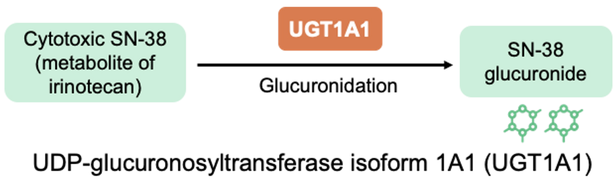
Racial Differences in *UGT1A1* Allele Frequencies and its Potential Impact in Pharmacogenetic Testing for Cancer Chemotherapy Drugs

Kyung Hyun Choi[1], Gene Selkov[1], Robert Huether[1], Brooke Rhead[1], and Francisco M. De La Vega[1]

[1]Tempus Labs, Chicago, IL

INTRODUCTION

UGT1A1 plays a crucial role in bilirubin metabolism and detoxification of many drugs such as irinotecan and atazanavir.



UDP-glucuronosyltransferase isoform 1A1 (UGT1A1)

Most *UGT1A1* pharmacogenetic (PGx) tests evaluate the risk of adverse events associated with irinotecan by genotyping a promoter TA-repeat polymorphism *28([TA]7) and*6, an exonic SNV (1).

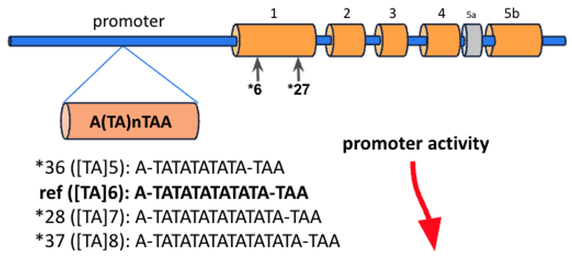


Figure 1. TA-repeat polymorphism in *UGT1A1* promoter

UGT1A1 polymorphism	Function/Characteristics
TA repeats (A[TA]nTAA, n = 5~8)	Inversely correlated with the transcriptional activity of the promoter
*36 ([TA]5)	Increased promoter activity Clinical significance not known
ref([TA]6)	Reference allele
*28 ([TA]7)	Reduced promoter activity Associated with elevated bilirubin
*37([TA]8)	Reduced promoter activity Clinical significance not known
*6 and *27	Nonsynonymous SNVs Associated with elevated bilirubin

Table 1. Clinically significant *UGT1A1* polymorphism (2)

We report a comprehensive analysis of *UGT1A1* TA-repeat and exonic variants from 14,569 de-identified patients. This study employed innovative methodologies such as:

- a clinical-grade NGS panel
- a novel TA-repeat caller algorithm
- computational phase analysis of compound heterozygotes
- imputation of race/ethnicity based on ancestry informative markers of the subjects

METHODS

- Genomic DNA from the normal tissues of 14,569 de-identified patients were sequenced by the Tempus xT.v4 germline NGS panel, a 648-gene panel with an average coverage 150x (3).
- The sequences were mapped by Burrows-Wheeler Aligner and *UGT1A1* exonic variants were called by DeepVariant (4).
- TA repeat length was determined using a bespoke Bayesian repeat calling algorithm.
- Annotation and *in silico* functional analysis of *UGT1A1* variants were performed using ANNOVAR (5).
- Six race/ethnicity categories — Non-Hispanic (NH) Black, NH East Asian, Hispanic or Latino, NH South Asian, NH White, and Complex — were imputed from continental genetic ancestry derived from 654 ancestry informative markers (AIMs), as described previously (6).
- Computational phasing of *UGT1A1* compound heterozygotes was performed using Eagle v.24 with a reference panel of 1000 Genome phase 3 (7).

RESULTS

Imputation of race/ethnicity from genetic ancestry

Patients were assigned to one of 6 race/ethnicity groups based on genetic ancestry proportions estimated from 654 AIMs covered by the xT.v4 germline NGS panel.

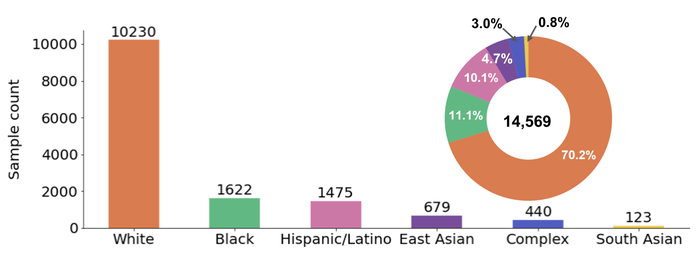


Figure 2. Imputed race/ethnicity of the cohort (n=14,569)

SIGNIFICANCE

- **We conducted a comprehensive analysis on the prevalence of *UGT1A1* genetic variants using a large diverse cohort.**
- **We identified significant differences in the prevalence of *UGT1A1* TA-repeat variants among five continental genetic ancestries.**
- **Determined the phase of *6 and *27, the 2 most prevalent *UGT1A1* SNVs with clinical significance, in relation to TA-repeat variants.**
- **Demonstrated the value of allele phasing in the assignment of *UGT1A1* metabolizer phenotype.**

The frequency of the TA repeat polymorphism varies significantly among different racial and ethnic groups.

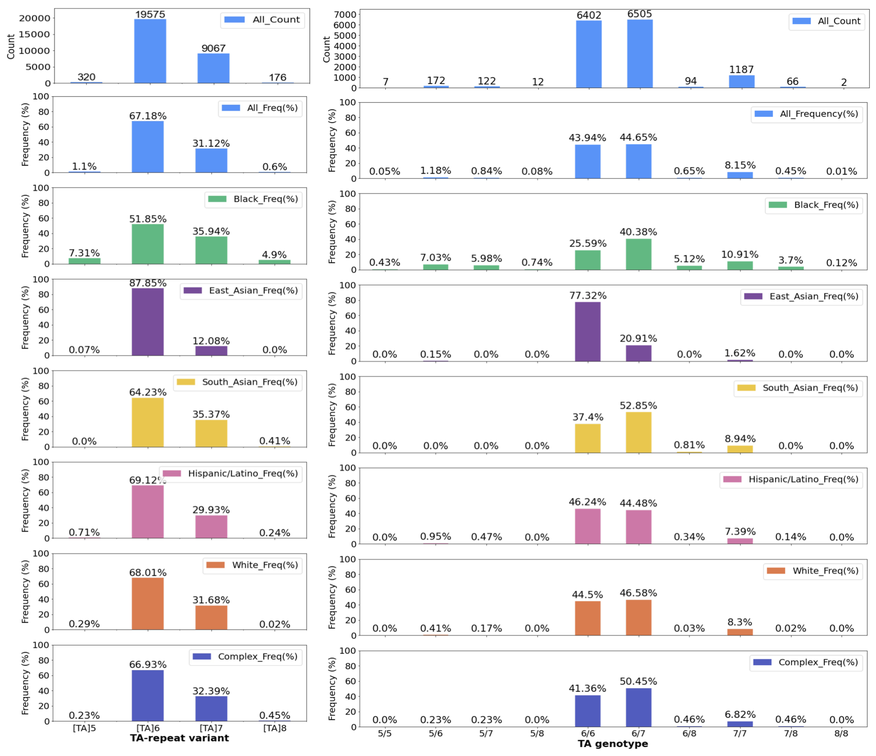


Figure 3. Frequency of TA-repeat variants and TA-repeat genotypes by genetic ancestry

- [TA]5 and [TA]8 have allele frequencies of 7.31% and 4.9%, respectively, in the NH Black imputed group but are mostly absent in other groups (<1%).
- The prevalence of genotypes containing [TA]5 or [TA]8 alleles is 23.12% in the NH Black imputed group and <2% in other groups.
- [TA]7 frequency is 12.08% in the NH East Asian imputed group, which is 2.5 times lower than the cohort's overall [TA]7 frequency.

More people carry TA-repeat genotypes with altered enzyme activities than the reference genotype TA6/TA6, except East Asians.

- [TA]6/[TA]7 is the most prevalent TA-repeat genotype at 44.65% in the cohort, which is closely followed by [TA]6/[TA]6 at 43.94%.

Total 62 potentially deleterious exonic variants are identified.

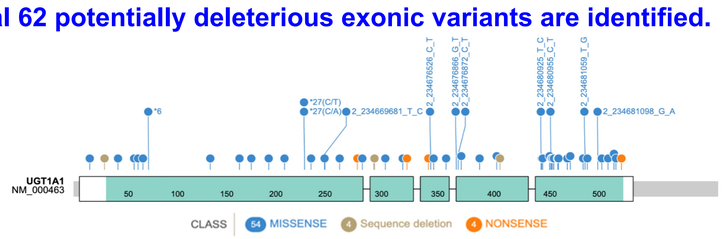


Figure 4. The location of 62 exonic variants with potential deleterious effect are marked on *UGT1A1* exons. The 10 most prevalent variants are noted.

- 121 unique exonic variants are identified from the cohort.
- 62 exonic variants are predicted to change the *UGT1A1* enzyme activity by *in silico* functional analysis.
  - 58 SNVs (DANN >= 0.96)
  - 4 deletion variants
- *6 is the most prevalent SNV with clinical significance, but mostly found in the NH East Asian imputed category (16.0%).
- *27(C/A) is exclusively found in NH East Asian (3.8%).
- Other SNVs with potential clinical significance is rare (<0.1%)

The prevalence of *UGT1A1* compound heterozygotes of altered enzyme activities are over 6% in East and South Asians.

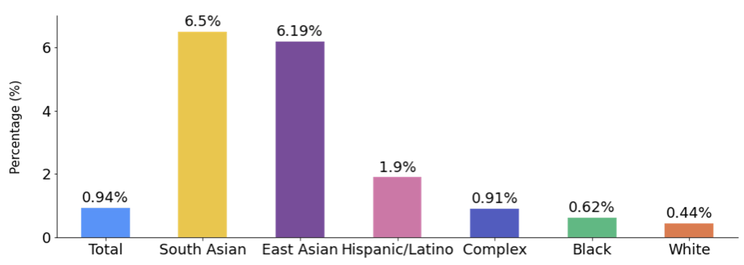


Figure 5. Frequency of *UGT1A1* compound heterozygotes with altered enzyme activities by genetic ancestry

- Total 137 cases of *UGT1A1* compound heterozygotes are identified (0.94%).
- 98.8% (134 cases) of the compound heterozygotes carry a TA-repeat variant.
- Due to the high allele frequencies of *6 and *27(C/A), compound heterozygote is highly prevalent in East and South Asians.

Phase status of the variant alleles in a compound heterozygote may change the metabolizer phenotype.

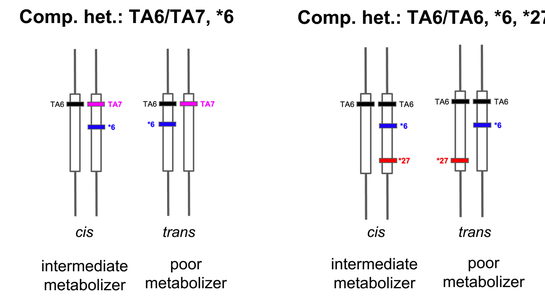


Figure 6. Phase status of hypothetical UGT1A1 compound heterozygotes and their metabolizer phenotypes based on CPIC guideline (8).

All *6's are in *trans* phase and all *27's in *cis* phase in relation to [TA]7 variant

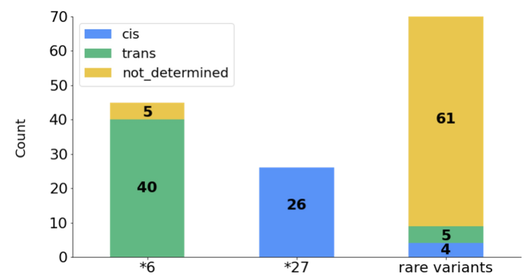


Figure 7. Result of computational phase analysis of 137 compound heterozygotes in relation to TA-repeat variant

- *27(C/A) is in complete LD with [TA]7 (D' =1, r^2= 0.005 in 1KG phase 3).
- *6 and *27 are always in *trans* phase to each other
- The phase of the compound heterozygotes with the rare variants are mostly not determined as the allele frequencies of the variants are too low in the 1KG reference panel.

Assignment of *UGT1A1* metabolizer phenotype with the phased genotype data

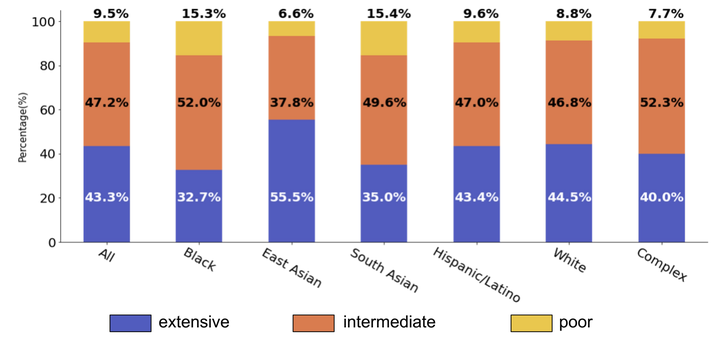


Figure 8. *UGT1A1* metabolizer phenotypes of the cohort.

- The assignment of metabolizer phenotype is done with [TA]6, [TA]7, [TA]8, phased *6 and *27 variants, and unphased SNVs with potential deleterious effect following CPIC *UGT1A1* metabolizer assignment scheme (8).
- [TA]5 is considered as a variant of normal function as its clinical effect is not determined yet.
- 2.8% of East Asian is reassigned from a poor metabolizer to an intermediate metabolizer based on the phase analysis.
- The haplotype *27-[TA]7 is considered as an allele of a reduced function, but its clinical effect is not known.

REFERENCES

1. Karas S and Innocenti F. All You Need to Know About UGT1A1 Genetic Testing for Patients Treated With Irinotecan: A Practitioner-Friendly Guide. JCO Oncol Pract. 2022 ;18(4):270-277.
2. Fujiwara Y and Minami H. An overview of the recent progress in irinotecan pharmacogenetics. Pharmacogenomics 2010;11(3):391–406
3. Beaubier et al., Clinical validation of the Tempus xT next-generation targeted oncology sequencing assay. Oncotarget. 2019;22;10(24):2384-2396.
4. Poplin R et al., A universal SNP and small-indel variant caller using deep neural networks. Nat Biotechnol. 2018;36(10):983-987
5. Wang K et al., ANNOVAR: Functional annotation of genetic variants from next-generation sequencing data. Nucleic Acids Res. 2010;38(16):e164.
6. Miyashita M. et al., Molecular profiling of a real-world breast cancer cohort with genetically inferred ancestries reveals actionable tumor biology differences between European ancestry and African ancestry patient populations. Breast Cancer Res. 2023;25;25(1):58
7. Das S et al., Next-generation genotype imputation service and methods. Nat Genet. 2016;48(10):1284-1287.
8. Gammal R. S. et al., Clinical Pharmacogenetics Implementation Consortium (CPIC) Guideline for UGT1A1 and Atazanavir Prescribing. Clin Pharmacol Ther. 2016;99(4):363-9.

Correspondence: Francisco M. De La Vega, Kyung H. Choi

Acknowledgement: We thank Amrita A. Iyer, Ph.D., for poster preparation & review