Supplementary Text

Novel variants in *NUDT15* and thiopurine intolerance in children with acute lymphoblastic leukemia from diverse ancestry

Takaya Moriyama[a], Yung-Li Yang[b], Rina Nishii[a, c], Hany Ariffin[d], Chengcheng Liu[a], Ting-Nien Lin[a], Wenjian Yang[a], Dong-Tsamn Lin[b], Chih-Hsiang Yu[e], Shirley Kham[f], Ching-Hon Pui[g], William E. Evans[a], Sima Jeha[g], Mary V. Relling[a], Allen Eng-Juh Yeoh[h, i] and Jun J. Yang[a]

[a]Department of Pharmaceutical Sciences, St. Jude Children's Research Hospital, Memphis, Tennessee, USA, [b]Department of Laboratory Medicine and Pediatrics, National Taiwan University Hospital, College of Medicine, National Taiwan University, Taipei, Taiwan, [c]Department of Pediatrics and Developmental Biology, Tokyo Medical and Dental University Graduate School of Medicine, Tokyo, Japan, [d]Paediatric Haematology-Oncology Unit, University of Malaya Medical Centre, Kuala Lumpur Malaysia, [e]Department of Clinical Laboratory Sciences and Medical Biotechnology, College of Medicine, National Taiwan University, Taipei, Taiwan, [f]National University Cancer Institute, National University Health System, Singapore, [g]Department of Oncology, St. Jude Children's Research Hospital, Memphis, Tennessee, USA, [h]Centre for Translational Research in Acute Leukaemia, Department of Paediatrics, Yong Loo Lin School of Medicine, National University of Singapore, Singapore, [i]Cancer Science Institute of Singapore, National University of Singapore, Singapore

NUDT15 and *TPMT* sequencing

Germline DNA was extracted from remission peripheral blood and *NUDT15* exons were sequenced following methods described previously.[1,2] cDNA synthesized from leukemia cell RNA was used to perform TA cloning (Invitrogen) in one Singaporean case carrying three *NUDT15* variants to resolve diplotype. *TPMT* variants (rs1800462, rs1800460, and rs1142345) were evaluated using Sanger sequencing.[3] Known *NUDT15* risk variants included p.R139C (rs116855232), p.R139H (rs147390019), p.V18I (rs186364861), and p.V18_V19insGV (rs554405994).

NUDT15 variant functional characterization

Novel *NUDT15* variants were cloned into the p.ColdII vector by site-directed mutagenesis (QuikChange II XL kit, Agilent Technologies). NUDT15 protein production, purification, diphosphatase activity, and thermostability measurements were performed following previously published methods.[2]

NUDT15 protein structure was drawn with PyMOL (Schrödinger, LLC) using the accession code 5LPG of PDB (http://www.rcsb.org/pdb/home/home.do).

Evaluation of erythrocyte thioguanine nucleotides (TGN) in patients

Whole blood of the St Jude TOTXVI case with the p.G17_V18del variant (Subject #5 in **Table 1**) was collected at week 7, 84 and 102 during his ALL maintenance therapy. Red blood cell processing and quantification of erythrocyte TGN by HPLC were performed according to previously published method[4,5]. This patient's red blood cell thiopurine metabolite levels were compared with that from patients with different genotype at the *NUDT15* p.R139C variant in the AALL03N1 cohort, as described previous[1,4,5] MP dosage two weeks prior to each red blood cell TGN measurement was used to standardize TGN levels.

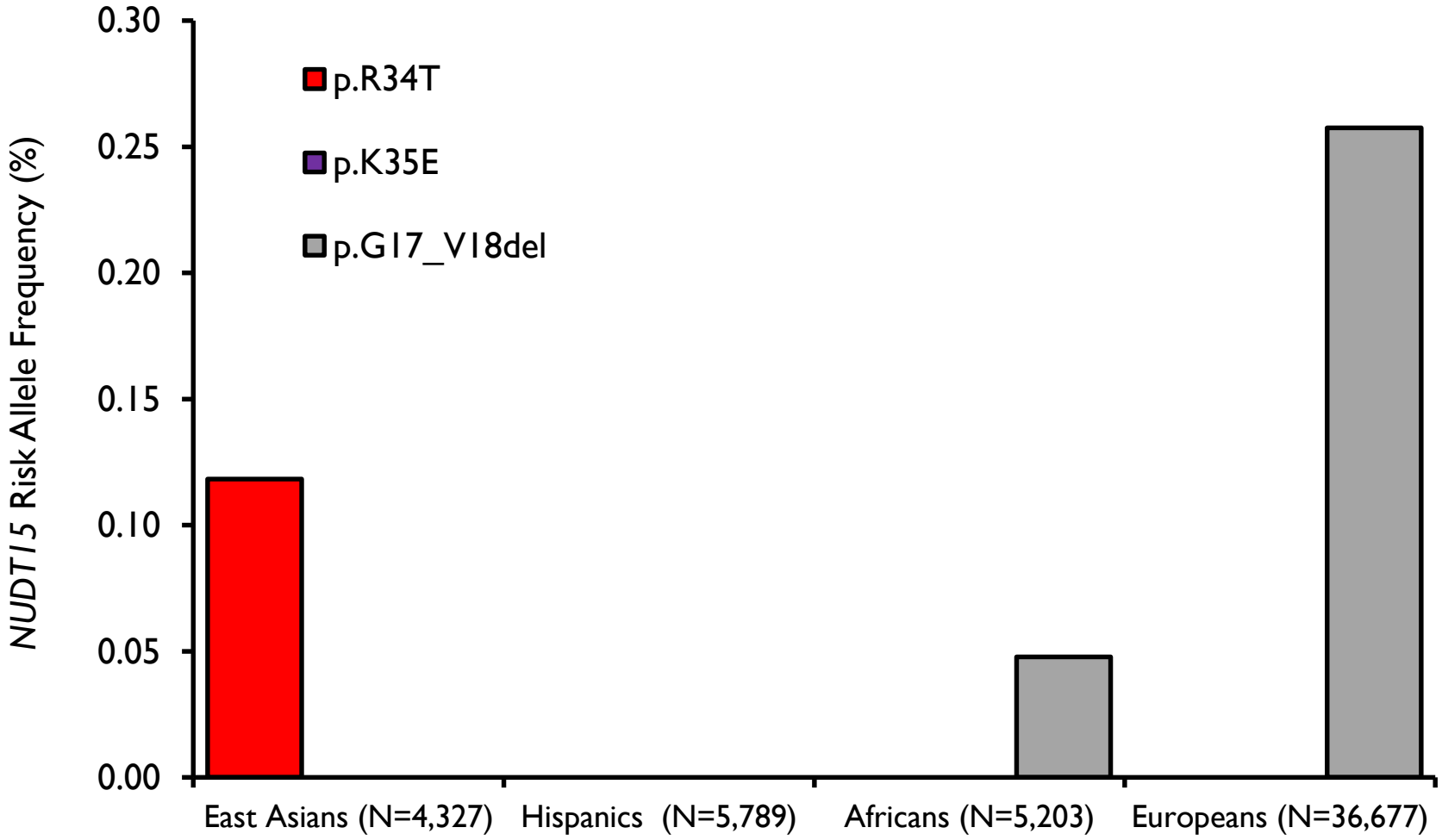
Effects on fitness of loss-of-function variants in *NUDT15*

There are currently no reports in the literature linking inherited *NUDT15* deficiency to health conditions in human. In fact, in a recent comprehensive analysis of fitness effects of loss-of-function variants in human genome[6], it was noted that there was no significant selection pressure against protein-truncating variants in *NUDT15* (S_{het} of 0.02), compared to for example cancer predisposition genes; and the distribution of *NUDT15* variants in large population dataset (i.e.,

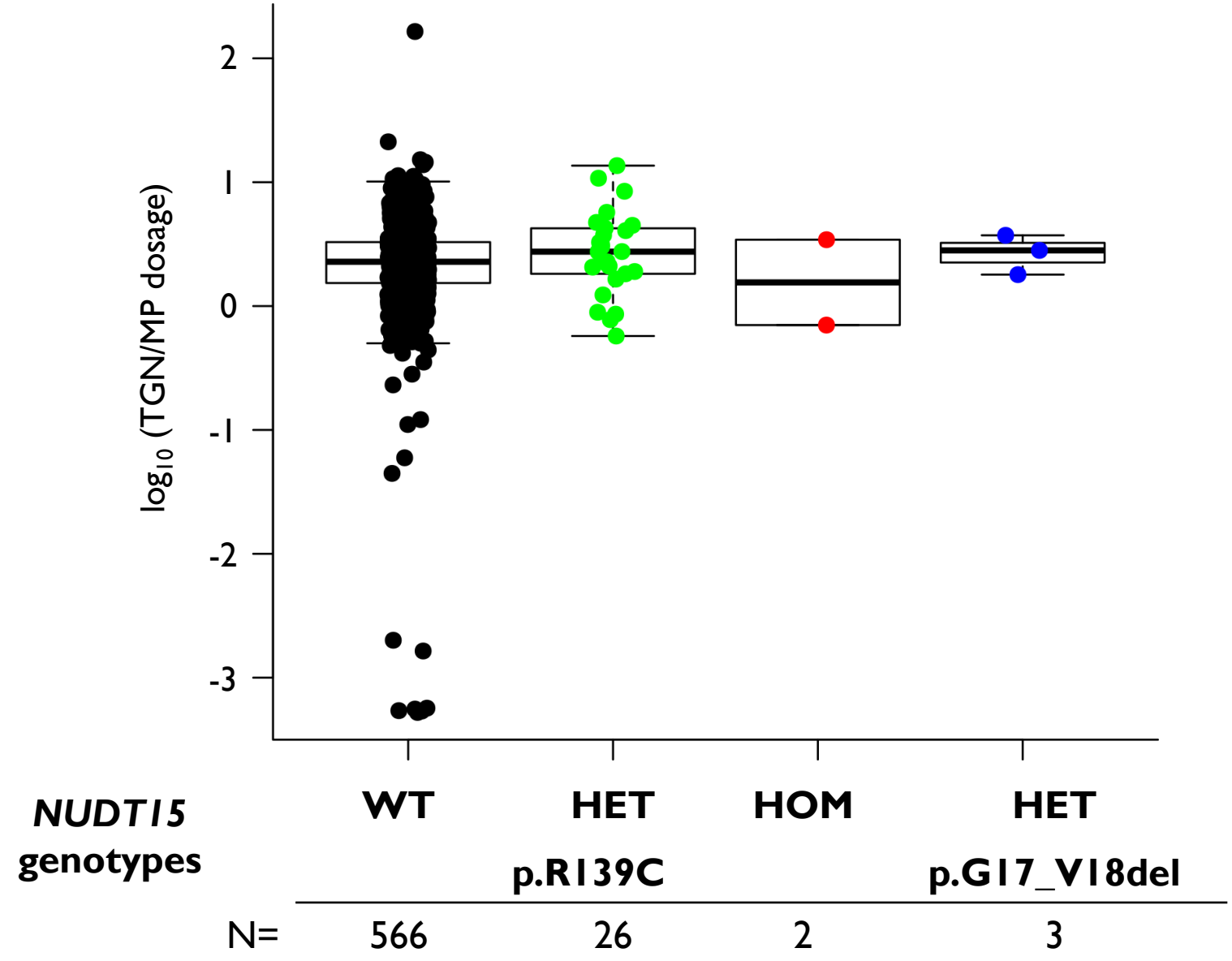
the ExAc cohort) does not support any association with loss of fitness under a dominant or recessive genetic model (P=0.26 and 0.74, respectively).

References

1. Yang JJ, Landier W, Yang W, et al. Inherited NUDT15 variant is a genetic determinant of mercaptopurine intolerance in children with acute lymphoblastic leukemia. *J Clin Oncol*. 2015;33(11):1235-1242.
2. Moriyama T, Nishii R, Perez-Andreu V, et al. NUDT15 polymorphisms alter thiopurine metabolism and hematopoietic toxicity. *Nat Genet*. 2016;48(4):367-373.
3. Yang JJ, Bhojwani D. Thiopurine S-methyltransferase pharmacogenetics in childhood acute lymphoblastic leukemia. *Methods Mol Biol*. 2013;999:273-284.
4. Bhatia S, Landier W, Hageman L, et al. 6MP adherence in a multiracial cohort of children with acute lymphoblastic leukemia: a Children's Oncology Group study. *Blood*. 2014;124(15):2345-2353.
5. Bhatia S, Landier W, Shangguan M, et al. Nonadherence to oral mercaptopurine and risk of relapse in Hispanic and non-Hispanic white children with acute lymphoblastic leukemia: a report from the children's oncology group. *J Clin Oncol*. 2012;30(17):2094-2101.
6. Cassa CA, Weghorn D, Balick DJ, et al. Estimating the selective effects of heterozygous protein-truncating variants from human exome data. *Nat Genet*. 2017;49(5):806-810.



Supplementary Figure 1. Frequency of novel *NUDT15* risk alleles by race and ethnicity Risk allele frequency was determined on the basis of the Board Institute Exome Aggregation Consortium (ExAc) database of 60,706 subjects with whole-exome seq.



Supplementary Figure 2 Comparison of erythrocyte TGN levels during MP treatment among patients with different *NUDT15* genotypes. The ratio of TGN to tolerated MP dosage was plotted against *NUDT15* genotype. The first 3 columns from the left represent data from patients with different genotypes at the p.R139C variant in the AALL03N1 cohort as described previously (e.g., J Clin Oncol 2012 30:2094). The column on the far right describes TGN levels at 3 time points during maintenance therapy of Subject #5 on the St Jude Total Therapy XVI with the p.G17_V18del variant.

novel NUDT15 **variants identified in the Broad Institute Exome Aggregation Consortium (ExAc) database**

Chrom	Position	rsID	Annotation	Protein Consequence	Alleles		Transcript Consequence	Minor Allele Frequency (%)					
					Reference	Alternative		All (N=60,706)	Africans (N=5,203)	Europeans (N=36,677)	East Asians (N=4,327)	South Asians (N=8,256)	Hispanics (N=5,789)
13	48037847	rs766023281	missense	p.Arg34Thr	R	T	c.101G>C	0.00825196	0	0	0.11829653	0	0
13	48037849	NA	missense	p.Lys35Glu	K	E	c.103A>G	NA	NA	NA	NA	NA	NA
13	48037783-48037788	rs746071566	inframe deletion	p.Gly17_Val18del	GV	del	c.37_42delGGAGTC	0.159686808	0.04784689	0.257506559	0	0.074595055	0