Supplementary Appendix

Supplement to: Treviño LR, Shimasaki N, Yang W, Panetta, et al. Germline genetic variation in an organic anion transporter polypeptide associated with methotrexate pharmacokinetics and clinical effects.

Methods

Genotyping calls and Imputations: Genotype calls were made using the Bayesian Robust Linear Multichip with Mahalanobis Distance (BRLMM) algorithm**.** Associations between germline SNP genotypes and methotrexate clearance were evaluated using a linear regression model, treating genotype as a numeric variable (0=AA, 1=AB, and 2=BB). Missing genotypes were imputed using linkage disequilibrium (LD) between missing and typed SNPs in the appropriate race/ancestry groups for each patient, with LD patterns established using the HapMap data sets release 22, when LD was informative. However, if the imputed genotype was discordant with the direction of the raw intensity signals (e.g. if the raw data indicated AA or AB but not BB genotype, but LD patterns indicated BB genotype), then the genotype was not imputed but left as missing. If LD information was unavailable for a SNP with missing genotype, an intermediate value (0 to 1 or 1 to 2) for genotype was assigned based on the actual signal intensity distribution observed from the chip. We attempted imputation for all SNPs on the 500K SNPCHIP. In total, 0.91% of total genotypes involving 58% of SNPs were successfully imputed, i.e. initial "NoCall" were imputed as "AA", "AB" or "BB" genotypes. Only 0.008% of all genotypes were left missing because of discordance between imputed genotypes and raw intensity signals. For all of the top 10 *SLCO1B1* SNPs listed in Table 1, the final call rate was greater than 98%. None of the *SLCO1B1* SNPs had genotypes that were left missing because of discordance between imputed genotypes and raw intensity signal values.

Gene-level Analysis: We performed a principal component analysis (PCA) using observed genotypes from all SNPs within the gene to account for the linkage disequilibrium among the SNPs. These derived genetic components from the PCA were uncorrelated with one another. We selected the top components (Supplemental table 11) that accounted for at least 95% of the genetic variation to be included in a multivariate analysis. The top two SNPs in the gene-level analysis, with the highest absolute value, were the same top-ranked SNPs in the single-SNP analysis. To test whether SNPs within a gene predict methotrexate clearance, we built multiple linear regression models using the derived genetic components as independent variables, and computed the r-squared test statistic as well as the Akaike's information criterion (AIC) for each gene. The AIC is a measure of the goodness-of-fit but also takes account of the number of predictors included in the model. To assess the overall significance of the gene, we performed 1×10^6 permutations. In each permutation, we randomly assigned methotrexate clearance values to patients and let the genetic components remain constant as observed. Using the permuted methotrexate clearance values and the observed genetic components, we performed the same multiple linear regression model and computed the r-squared statistic. The proportion of permutations with r-squared values larger than or equal to the observed was assigned as the significance of the gene. By the above analysis, each gene had only one P value and AIC estimate (Supplemental table 7) thus avoiding the multiple testing caused by multiple SNPs within the gene.

Ancestry and Race determination: Genetically-determined race was assessed two ways: (1) using the percentage of ancestry (European, African, and Asian) for each patient, and (2) using the Eigenstrat method of principal component analysis. To estimate the percentage of ancestry for each patient, approximately 100,000 SNPs from the total possible 398,699 evaluable SNPs with a 100% genotyping call rate (which were also genotyped in the HapMap panels of African, Asian and European origin) were selected. We randomly selected 100 from the pool of 100,000 SNPs, and built a prediction model using a linear discriminant analysis based on the HapMap genotype data that distinguished individuals of African, Asian and European ancestry. We predicted the probability of a patient being of European descent, African descent and Asian descent based on the model. We repeated the above procedure 1000 times, and computed the average probability as the estimate of percent of European, African, or Asian ancestry for each patient; each patient had an estimated percentage of European ancestry, percentage of African ancestry, and percentage of Asian ancestry. We found that 5 patients had ancestral genotypes consistent with at least 90% Asian ancestry, 73 patients with at least 90% African-American ancestry, and 315 patients with at least 90% Caucasian ancestry. The remaining 41 patients were classified as "other." There was good concordance (94%) between self-reported race and race classified by genotype data on the 500K SNPCHIP. We also used a principal component analysis for assessing racial composition based on the Eigenstrat method[1].

Standardization of non-genetic factors: Hydration, alkalinization, dosing, and monitoring were standardized because all patients were prospectively studied on single-institution research protocols that used extensive infrastructure and protocol-specific standard orders (Supplemental table 2). Each course of high-dose methotrexate was accompanied by 2 pages of single-spaced, detailed instructions that included hydration fluids, IV rates, alkalinization with sodium bicarbonate, urine pH monitoring, instructions for further alkalinization, clinical chemistries, avoidance of interacting drugs, plasma monitoring of methotrexate concentrations, antiemetics, contingency plans for vomiting, low urine output or low urine pH and leucovorin rescue. All toxicity was monitored and reviewed at biweekly program meetings. All patients and families were interviewed by a clinical pharmacist and non-protocol supportive care was changed, if needed, to minimize drug interactions or other clinical features that might affect methotrexate clearance and risk for delayed excretion.

Prospective grading of toxicity: Toxicity was graded using the NCI CTEP guidelines (http://ctep.cancer.gov). All data including grade 3 and 4 toxicities were carefully summarized and reviewed by the principal investigator prior to submission to an external DSMB. These data were retrieved from established institutional databases and discussed in detail by both principal and co-investigators at protocol meetings held every two weeks. Although graded prospectively, the relationship between toxicities and SNP genotypes was

evaluated retrospectively, after results of the initial genotype vs. methotrexate clearance analysis were known.

Supplemental table 1. Number of patients per treatment protocol; all have both SNP genotype and methotrexate clearance data available.

Protocol	Methotrexate dose	No. of Courses	Discovery Cohort (n=434)		Validation Cohort (n = 206)	
			SNPChip data (n= 434)	Subset with T521C genotyping (n=387)	SNPChip data (n=206)	Subset with T521C genotyping (n=102)
Total XIIIB 1994 - 1998	2 g/m^2/2 hours	10	213	174	0	0
Total XV[a] low risk 2000 – 2007	2.5 g/m^2/24 hours	4	114	110	97	50
Total XV[a] standard/high risk 2000 – 2007	5 g/m^2/24 hours	4	107	103	109	52

[a]For Total XV, patients on the low risk arm and the standard/high risk arm had dosages individualized to achieve a steady state concentration of 33 μM or 65 μM respectively.

Supplemental table 2. Details of methotrexate administration, hydration and alkalinization.

	Total XIIIB	Total XV low risk	Total XV standard/high risk
Methotrexate dose	2000 mg/m^2 in 100 ml D_5W IV over 2 hr	*250 mg/m^2 over 1 hr then 2250 mg/m^2/23 hr	**500 mg/m^2 over 1 hr then 4500 mg/m^2/23 hr
Hydration duration	From at least 1 hr before methotrexate to ≥ 5 hr after methotrexate	From at least 2 hr before methotrexate to ≥ 42 hr after Methotrexate	From at least 2 hr before methotrexate to ≥ 42 hr after methotrexate
Hydration rate	200 ml/m^2/hr plus bolus of 12 mEq/m^2 $NaHCO_3$	100 ml/m^2/hr if started night before methotrexate; if not, 200 ml/m^2/hr x 2 hr + 25 mEq/m^2 $NaHCO_3$, then 100 ml/m^2/hr.	125ml/m^2/hr if started night before methotrexate; if not, 200 ml/m^2/hr x 2 hr+ 25 mEq/m^2 $NaHCO_3$, then 125 ml/m^2/hr
Hydration fluid	D_5W ¼ NS $NaHCO_3$ 40 mEq/L	D_5W + KCl 20 mEq/L + $NaHCO_3$ 40 mEq/L	D_5W + KCl 20 mEq/L + $NaHCO_3$ 40 mEq/L
Urine pH = 6.0 Urine pH < 6.0	Give $NaHCO_3$ 12.5 mEq/m^2 Give $NaHCO_3$ 25 mEq/m^2	Give $NaHCO_3$ 12.5 mEq/m^2 Give $NaHCO_3$ 25 mEq/m^2	Give $NaHCO_3$ 12.5 mEq/m^2 Give $NaHCO_3$ 25 mEq/m^2
Plasma for methotrexate at:	Pre, 1,6,21 & 44 hr from start	Pre, 6, 23 & 42 hr from start	Pre, 6, 23 & 42 hr from start
Baseline leucovorin (dose is increased from baseline for elevated plasma methotrexate)	10 mg/m^2 q6 hr x 5 to start at 44 hr	10 mg/m^2/dose q6 hr x 5 to start at 42 hr	15 mg/m^2/dose q6 hr x 5 start at 42 hr

**doses individualized to achieve a plasma steady state methotrexate concentration of 33µM*
*** doses individualized to achieve a plasma steady state methotrexate concentration of 65µM*

Supplemental table 3. SNPs associated with average methotrexate clearance in the discovery set of 434 children with ALL.

dbSNP ID	Chr.	Alleles (A/B)	P value[a]	Eigenstrat P value[b]	MAF	Gene symbol
rs11045879	chr12	*T/C	$1.7X\ 10^{-10}$	$1.6\ X\ 10^{-10}$	0.16	*SLCO1B1*
rs4149081	chr12	*G/A	$1.7X\ 10^{-9}$	$4.0\ X\ 10^{-9}$	0.16	*SLCO1B1*
rs16923647	chr12	T/C*	$3.64X\ 10^{-6}$	$3.64\ X\ 10^{-6}$	0.13	*SLCO1A2*
rs12576286	chr11	A/T*	$1.52X\ 10^{-5}$	$7.16\ X\ 10^{-6}$	0.02	*SOX6*
rs3737416	chr21	*C/T	$1.76X\ 10^{-5}$	$1.75\ X\ 10^{-5}$	0.05	*APP*
rs8065836	chr17	A/G*	$2.27X\ 10^{-5}$	$1.36\ X\ 10^{-5}$	0.45	*TNFRSF13B*
rs17781156	chr14	T/C*	$2.28X\ 10^{-5}$	$2.83\ X\ 10^{-5}$	0.5	
rs4338756	chr15	A/C*	$2.57X\ 10^{-5}$	$7.97\ X\ 10^{-6}$	0.14	
rs7912575	chr10	G/A*	$2.64X\ 10^{-5}$	$5.32\ X\ 10^{-6}$	0.02	*NRG3*
rs468522	chr21	*G/A	$2.80X\ 10^{-5}$	$1.92\ X\ 10^{-5}$	0.16	
rs12523702	chr6	C/T*	$3.81X\ 10^{-5}$	$2.79\ X\ 10^{-5}$	0.37	
rs869532	chr8	*A/C	$3.85X\ 10^{-5}$	$2.01\ X\ 10^{-5}$	0.04	
rs128647	chr21	*G/A	$3.91X\ 10^{-5}$	$3.87\ X\ 10^{-5}$	0.04	*APP*
rs12894524	chr14	*T/G	$3.92X\ 10^{-5}$	$1.31\ X\ 10^{-4}$	0.38	*MYH7*
rs7040439	chr9	T/A*	$4.06X\ 10^{-5}$	$3.27\ X\ 10^{-5}$	0.05	
rs468969	chr21	*A/G	$4.22X\ 10^{-5}$	$2.55\ X\ 10^{-5}$	0.16	
rs4324007	chr13	*A/T	$4.77X\ 10^{-5}$	$6.22\ X\ 10^{-5}$	0.31	*GPC5*
rs2881582	chr15	*T/C	$5.97X\ 10^{-5}$	$1.91\ X\ 10^{-5}$	0.18	*AGBL1*
rs17089780	chr4	*C/T	$5.98X\ 10^{-5}$	$3.47\ X\ 10^{-4}$	0.02	
rs6448167	chr4	G/C*	$6.00X\ 10^{-5}$	$2.22\ X\ 10^{-5}$	0.05	
rs17502069	chr18	A/T*	$6.27X\ 10^{-5}$	$4.25\ X\ 10^{-5}$	0.14	*L3MBTL4*
rs1863603	chr2	G/A*	$6.92X\ 10^{-5}$	$6.26\ X\ 10^{-5}$	0.21	
rs7704890	chr5	T/C*	$7.06X\ 10^{-5}$	$5.24\ X\ 10^{-5}$	0.35	*MAST4*
rs16924862	chr11	*A/G	$7.18X\ 10^{-5}$	$4.83\ X\ 10^{-5}$	0.02	*SOX6*
rs16829700	chr3	*G/A	$8.43X\ 10^{-5}$	$1.21\ X\ 10^{-4}$	0.06	*CDGAP*
rs960390	chr21	T/G*	$9.12X\ 10^{-5}$	$4.0\ X\ 10^{-4}$	0.34	
rs7568533	chr2	A/C*	$9.24X\ 10^{-5}$	$9.65\ X\ 10^{-5}$	0.21	
rs1887365	chr4	*C/G	$9.28X\ 10^{-5}$	$2.23\ X\ 10^{-4}$	0.17	*ZFYVE28*
rs2754163	chr14	T/C*	$9.31X\ 10^{-5}$	$1.12\ X\ 10^{-4}$	0.32	*MYH7*
rs4129601	chr11	G/A*	$9.63X\ 10^{-5}$	$5.44\ X\ 10^{-5}$	0.03	
rs529270	chr11	*T/G	$9.75X\ 10^{-5}$	$5.65\ X\ 10^{-5}$	0.06	*STX3*
rs3758729	chr11	*C/T	$9.76X\ 10^{-5}$	$7.64\ X\ 10^{-5}$	0.46	*PRDM11*
rs3829767	chr14	*G/A	$1.01X\ 10^{-5}$	$7.33\ X\ 10^{-5}$	0.16	*SYNE2*
rs6966641	chr7	G/A*	$1.03X\ 10^{-5}$	$8.17\ X\ 10^{-5}$	0.43	
rs4722283	chr7	*G/C	$1.03X\ 10^{-4}$	$8.45\ X\ 10^{-5}$	0.43	
rs2057200	chr14	*G/A	$1.04X\ 10^{-4}$	$1.60\ X\ 10^{-4}$	0.05	
rs9845965	chr3	*T/G	$1.07X\ 10^{-4}$	$1.51\ X\ 10^{-4}$	0.07	*CDGAP*
rs4020077	chr2	*A/G	$1.08X\ 10^{-4}$	$1.04\ X\ 10^{-4}$	0.08	
rs17196492	chr13	T/C*	$1.09X\ 10^{-4}$	$3.65\ X\ 10^{-4}$	0.07	
rs1317057	chr7	A/G*	$1.10X\ 10^{-4}$	$4.88\ X\ 10^{-5}$	0.24	*EXOC4*
rs309043	chr12	*T/G	$1.13X\ 10^{-4}$	$1.73\ X\ 10^{-4}$	0.35	

rs17564624	chr5	G/A*	$1.17X\ 10^{-4}$	$1.87\ X\ 10^{-4}$	0.04	*KCNIP1*
rs11645471	chr16	*A/T	$1.19X\ 10^{-4}$	$3.01\ X\ 10^{-4}$	0.03	*NKD1*
rs11138867	chr9	G/A*	$1.20X\ 10^{-4}$	$9.3\ X\ 10^{-5}$	0.04	
rs2824008	chr21	T/C*	$1.25X\ 10^{-4}$	$3.17\ X\ 10^{-4}$	0.34	
rs1255971	chr14	T/G*	$1.27X\ 10^{-4}$	$9.2\ X\ 10^{-5}$	0.15	*SYNE2*
rs7036018	chr9	T/C*	$1.28X\ 10^{-4}$	$1.79\ X\ 10^{-4}$	0.14	*SLC24A2*
rs10743623	chr12	A/G*	$1.29X\ 10^{-4}$	$1.84\ X\ 10^{-4}$	0.4	
rs10740455	chr10	*T/A	$1.31X\ 10^{-4}$	$1.26\ X\ 10^{-4}$	0.11	*C10orf11*
rs10841753	chr12	T/C*	$1.34X\ 10^{-4}$	$1.52\ X\ 10^{-4}$	0.19	*SLCO1B1*

Shown is the dbSNP ID, chromosome, SNP alleles, P value of association between the listed SNP and methotrexate clearance, the minor allele frequency in the discovery set (n=434) and the gene symbol annotated to the listed SNP.

*Alleles denoted with an asterisk have higher methotrexate clearance.

[a] The P value shown represents the evidence of association between the polymorphisms and methotrexate clearance, adjusting for age, treatment regimen and race assessed as the percentage of an individual's ancestry (percentage European, percentage African and percentage Asian) as covariates.

[b] The P value shown represents the evidence of association between the polymorphisms and methotrexate clearance, adjusting for age, treatment regimen and race assessed using principal components, as described by Eigentstrat.[1]

Supplemental table 4. SNPs associated with methotrexate clearance (in a linear mixed effects model, retaining clearance data for each course) in the discovery cohort of 434 children with ALL.

dbSNP ID	Chr.	Alleles (A/B)	P value	MAF	Gene symbol
rs11045879	chr12	*T/C	$9.55X\ 10^{-11}$	0.16	*SLCO1B1*
rs4149081	chr12	*G/A	$6.80X\ 10^{-9}$	0.16	*SLCO1B1*
rs16923647	chr12	T/C*	$1.30X\ 10^{-6}$	0.13	*SLCO1A2*
rs3737416	chr21	*C/T	$9.05X\ 10^{-6}$	0.05	*APP*
rs12576286	chr11	A/T*	$9.06X\ 10^{-6}$	0.02	*SOX6*
rs7912575	chr10	G/A*	$1.72X\ 10^{-5}$	0.02	*NRG3*
rs17781156	chr14	T/C*	$1.84X\ 10^{-5}$	0.5	
rs4324007	chr13	*A/T	$2.13X\ 10^{-5}$	0.31	*GPC5*
rs128647	chr21	*G/A	$2.14X\ 10^{-5}$	0.04	*APP*
rs468522	chr21	*G/A	$2.46X\ 10^{-5}$	0.16	
rs12894524	chr14	*T/G	$3.04X\ 10^{-5}$	0.38	*MYH7*
rs869532	chr8	*A/C	$3.13X\ 10^{-5}$	0.04	
rs4338756	chr15	A/C*	$3.22X\ 10^{-5}$	0.14	
rs8065836	chr17	A/G*	$3.58X\ 10^{-5}$	0.45	*TNFRSF13B*
rs468969	chr21	*A/G	$3.77X\ 10^{-5}$	0.16	
rs6448167	chr4	G/C*	$3.84X\ 10^{-5}$	0.05	
rs2881582	chr15	*T/C	$3.98X\ 10^{-5}$	0.18	*AGBL1*
rs1317057	chr7	A/G*	$4.04X\ 10^{-5}$	0.24	*EXOC4*
rs16924862	chr11	*A/G	$4.05X\ 10^{-5}$	0.02	*SOX6*
rs4020077	chr2	*A/G	$4.07X\ 10^{-5}$	0.08	
rs16829700	chr3	*G/A	$4.43X\ 10^{-5}$	0.06	*CDGAP*
rs878815	chr14	*G/A	$5.42X\ 10^{-5}$	0.4	
rs9845965	chr3	T/G*	$5.60X\ 10^{-5}$	0.07	*CDGAP*
rs17089780	chr4	*C/T	$5.71X\ 10^{-5}$	0.02	
rs12997105	chr2	G/C*	$5.71X\ 10^{-5}$	0.08	
rs17502069	chr18	A/T*	$5.95X\ 10^{-5}$	0.14	*L3MBTL4*
rs4832104	chr2	*C/A	$6.29X\ 10^{-5}$	0.08	
rs7040439	chr9	T/A*	$6.48X\ 10^{-5}$	0.05	
rs7704890	chr5	T/C*	$6.70X\ 10^{-5}$	0.35	*MAST4*
rs1887365	chr4	*C/G	$6.70X\ 10^{-5}$	0.17	*ZFYVE28*
rs4129601	chr11	G/A*	$6.87X\ 10^{-5}$	0.03	
rs12523702	chr6	C/T*	$6.94X\ 10^{-5}$	0.37	
rs7036018	chr9	*T/C	$8.10X\ 10^{-5}$	0.14	*SLC24A2*
rs10841753	chr12	T/C*	$8.14X\ 10^{-5}$	0.19	*SLCO1B1*
rs3829767	chr14	*G/A	$8.46X\ 10^{-5}$	0.16	*SYNE2*
rs10743623	chr12	A/G*	$8.47X\ 10^{-5}$	0.4	
rs309043	chr12	*T/G	$8.87X\ 10^{-5}$	0.35	
rs1625289	chr21	*G/C	$9.10X\ 10^{-5}$	0.05	*APP*
rs17564624	chr5	G/A*	$9.11X\ 10^{-5}$	0.04	*KCNIP1*
rs6847225	chr4	T/A*	$9.26X\ 10^{-5}$	0.47	*UCHL1*
rs7746263	chr6	G/C*	$9.29X\ 10^{-5}$	0.03	
rs10210186	chr2	T/C*	$9.50X\ 10^{-5}$	0.08	
rs2754163	chr14	T/C*	$9.63X\ 10^{-5}$	0.32	*MYH7*
rs184649	chr18	G/A*	$9.82X\ 10^{-5}$	0.19	

rs6027462	chr20	A/G*	$1.01X\ 10^{-4}$	0.22
rs17615623	chr19	*C/A	$1.06X\ 10^{-4}$	0.15
rs6966641	chr7	G/A*	$1.11X\ 10^{-4}$	0.43
rs1990763	chr19	*A/G	$1.11X\ 10^{-4}$	0.16
rs11126970	chr2	*G/A	$1.11X\ 10^{-4}$	0.08
rs17125091	chr14	*A/T	$1.12X\ 10^{-4}$	0.21

Shown is the dbSNP ID, chromosome, SNP alleles, P value of association between the listed SNP and methotrexate clearance, the minor allele frequency in the discovery set (n=434) and the gene symbol annotated to the listed SNP.

*Alleles denoted with an asterisk have higher methotrexate clearance.

Supplemental table 5. SNPs associated with methotrexate clearance in the combined (discovery plus validation) cohort of 640 children with ALL.

dbSNP ID	Chr	Alleles (A/B)	P value	Eigenstrat P value	MAF	Gene symbol
rs11045879	chr12	T/C*	7.64 X 10-11	1.3 X 10-10	0.16	SLCO1B1
rs4149081	chr12	G/A*	1.47 X 10-9	2.2 X 10-9	0.16	SLCO1B1
rs11045818	chr12	A/G*	6.73 X 10-7	1.1 X 10-5	0.13	SLCO1B1
rs16923647	chr12	*T/C	1.59 X 10-6	7.8 X 10-5	0.10	SLCO1A2
rs10740455	chr10	T/A*	6.77 X 10-6	9.8 X 10-6	0.08	C10orf11
rs10841753	chr12	* T/C	6.93 X 10-6	1.2 X 10-5	0.19	SLCO1B1
rs16951021	chr16	C/T*	9.02 X 10-6	3.2 X 10-7	0.01	
rs6543610	chr2	G/C*	9.95 X 10-6	1.4 X 10-7	0.02	
rs16924862	chr11	*A/G	1.02 X 10-5	3.1 X 10-5	0.04	SOX6
rs7742587	chr6	T/C*	1.1 X 10-5	2.8 X 10-7	0.02	
rs4129601	chr11	G/A*	1.54 X 10-5	2.3 X 10-6	0.03	
rs6570465	chr6	T/C*	1.63 X 10-5	3.9 X 10-7	0.02	
rs12469413	chr2	*A/G	1.87 X 10-5	5.9 X 10-5	0.22	
rs2763993	chr6	T/C*	1.97 X 10-5	2.9 X 10-4	0.15	PACRG
rs6476012	chr9	G/A*	2.06 X 10-5	1.0 X 10-4	0.38	
rs495524	chr1	*T/A	2.17 X 10-5	1.5 X 10-6	0.42	
rs7312122	chr12	*G/A	2.28 X 10-5	8.4 X 10-6	0.16	
rs404801	chr3	A/C*	2.36 X 10-5	5.3 X 10-5	0.45	
rs903428	chr15	C/G*	2.46 X 10-5	1.9 X 10-5	0.33	
rs16842994	chr2	*C/G	2.55 X 10-5	1.6 X 10-6	0.02	
rs12576286	chr11	A/T*	2.67 X 10-5	3.5 X 10-4	0.03	SOX6
rs276000	chr12	*C/T	3.06 X 10-5	7.2 X 10-7	0.02	
rs564816	chr9	A/C*	3.07 X 10-5	2.7 X 10-7	0.04	GLIS3
rs4681523	chr3	*A/G	3.09 X 10-5	6 X 10-5	0.08	WWTR1
rs11045872	chr12	*A/G	3.34 X 10-5	2.6 X 10-4	0.15	SLCO1B1
rs7568533	chr2	*A/C	3.53 X 10-5	1.1 X 10-4	0.22	
rs1863603	chr2	*G/A	3.54 X 10-5	1.0 X 10-4	0.22	
rs17017982	chr4	*A/G	3.98 X 10-5	5.4 X 10-4	0.01	
rs17781156	chr14	*T/C	4.39 X 10-5	2.4 X 10-5	0.50	
rs350585	chr3	T/C*	4.55 X 10-5	1.2 X 10-4	0.45	
rs1396293	chr3	T/G*	4.59 X 10-5	2 X 10-5	0.24	
rs6873545	chr5	T/C*	4.69 X 10-5	9.9 X 10-4	0.30	GHR
rs10495048	chr1	*C/T	4.94 X 10-5	1.6 X 10-4	0.26	
rs7704890	chr5	*T/C	5.17 X 10-5	1.1 X 10-5	0.33	MAST4
rs4145130	chr6	T/A*	5.21 X 10-5	1.2 X 10-7	0.06	
rs4339926	chr1	A/G*	5.37 X 10-5	1.1 X 10-4	0.26	
rs7523455	chr1	G/A*	5.53 X 10-5	1.5 X 10-7	0.19	
rs468522	chr21	G/A*	5.63 X 10-5	1.5 X 10-4	0.14	
rs2479768	chr13	*G/A	5.72 X 10-5	5.5 X 10-6	0.03	KIAA0774
rs4387258	chr10	A/G*	5.77 X 10-5	3.5 X 10-6	0.14	
rs16862426	chr1	T/C*	5.88 X 10-5	8.5 X 10-8	0.03	
rs6448167	chr4	*G/C	6.04 X 10-5	0.008	0.07	
rs16904316	chr8	T/C*	6.13 X 10-5	1.6 X 10-6	0.03	
rs11166411	chr1	*C/T	6.23 X 10-5	3.8 X 10-5	0.11	LRRC39

rs4324007	chr13	A/T*	6.91 X 10-5	8.3 X 10-5	0.31	GPC5
rs8004448	chr14	*A/G	6.91 X 10-5	0.002	0.36	
rs17032807	chr4	T/C*	7.62 X 10-5	5.3 X 10-5	0.07	
rs7036018	chr9	T/C*	7.81 X 10-5	7.7 X 10-5	0.12	SLC24A2
rs12595731	chr15	G/A*	7.95 X 10-5	3.3 X 10-5	0.13	

Shown is the dbSNP ID, chromosome, SNP alleles, P value of association between the listed SNP and methotrexate clearance, the minor allele frequency in the combined patient cohort (n=640) and the gene symbol annotated to the listed SNP.

*Alleles denoted with an asterisk have higher methotrexate clearance.

[a] The P value shown represents the evidence of association between the polymorphisms and methotrexate clearance, adjusting for age, treatment regimen and race assessed as the percentage of an individual's ancestry (percentage European, percentage African and percentage Asian) as covariates.

[b] The P value shown represents the evidence of association between the polymorphisms and methotrexate clearance, adjusting for age, treatment regimen and race assessed using principal components, as described by Eigentstrat.[1]

Supplemental table 6. Results from a linear mixed effects model in the combined patient cohort (n=559 patients with creatinine measurements) including course number and creatinine as a covariates.

		Coefficient	P value	Percentage of interpatient variability	Percentage of intrapatient variability
Race [a]				2.10 %	0%
	European	*			
	African	11.0	2.5×10^{-5}		
	Asian	10.5	0.21		
Gender				2.7%	0%
	Female	*			
	Male	6.3	8.0×10^{-4}		
Creatinine [b] ***(mg/dl)***		-21.4	6.6×10^{-6}	2.4%	0.4%
Age (years)		-0.43	0.08	0.1%	0%
Treatment [c]				18.6%	0%
	Total XIIIB	*			
	Total XV S/HR	-19.0	3.4×10^{-10}		
	Total XV LR	-7.2	5.5×10^{-6}		
Courses				0%	6.0%
	Total XIIIB Course 1	*			
	Total XIIIB Course 2	-9.6	3.7×10^{-5}		
	Total XIIIB Course 3	7.1	0.003		
	Total XIIIB Course 4	9.0	1.0×10^{-4}		
	Total XIIIB Course 5	-2.6	0.26		
	Total XIIIB Course 6	-4.4	0.061		
	Total XIIIB Course 7	4.8	0.040		
	Total XIIIB Course 8	5.7	0.016		
	Total XIIIB Course 9	7.2	0.0023		
	Total XIIIB Course 10	2.5	0.30		
	Total XV Course 1	*			
	Total XV Course 2	-10.0	7.4×10^{-7}		
	Total XV Course 3	-10.3	1.0×10^{-7}		
	Total XV Course 4	-9.6	2.8×10^{-7}		
SLCO1B1 genotype				11.0%	0%
	rs4149081	-12.4	3.6×10^{-7}		
	rs17328763	7.6	0.061		
	rs10841753	6.8	0.018		
	rs2900476	-3.8	0.078		
	rs11045818	-8.6	0.065		

* Denotes the group used as reference in the multivariate analysis.

[a] Race was assessed as the percentage of an individual's ancestry (percentage European, percentage African and percentage Asian)

[b] The serum creatinine concentration (mg/dl) before each methotrexate course was used as a measure for renal function and included in the multivariate analysis.

[c] The coefficient estimates for the Treatment covariate compare the methotrexate clearance of the first course of the Total XIIIB protocol (reference) to the first methotrexate clearance of the first course of the Total XV Standard/High Risk arm or the first methotrexate clearance of the first course of the Total XV Low Risk arm.

Supplemental table 7. The *SLCO1B1* gene was most highly associated with methotrexate clearance in a gene-by gene analysis.

Gene Symbol	No. of SNPs in gene on SNP chip	No. of components used in the model	Discovery Set (n=434)			Validation Set (n=206)		
			r^2	AIC	P value	r^2	AIC	P value
SLCO1B1	27	11	0.134	3933.786	$< 1 \times 10^{-6}$	0.099	1805.7	0.024
LOC96597	3	2	0.048	3956.899	2.3×10^{-5}	0.005	1810.3	0.62
MYH6	6	4	0.060	3955.365	2.6×10^{-5}	0.015	1812.2	0.56
ZNF556	3	3	0.049	3958.512	8.1×10^{-5}	0.004	1810.4	0.65
TRBV6-6	3	1	0.035	3960.873	9.7×10^{-5}	<0.001	1809.3	0.98
SLCO1A2	25	11	0.082	3959.275	1.2×10^{-4}	0.081	1811.9	0.12
FOXN2	12	5	0.050	3962.080	5.2×10^{-4}	0.046	1807.6	0.10
CYB5A	22	8	0.061	3963.119	7.9×10^{-4}	0.040	1812.8	0.29
IFNA10	14	2	0.031	3964.659	0.0012	0.006	1810.1	0.56

Listed are the top nine genes that were significantly associated with methotrexate clearance in the global gene-level analysis performed in the discovery cohort (n=434) and validation cohort (n=206). Shown is the gene symbol, the number of SNPs for each gene that is represented on the mapping array, the number of components used in the multiple linear regression model, the estimated r-squared value, the Akaike information criterion (AIC) value, and the P value. The

SLCO1B1 locus had the lowest P value, the highest r-squared value and the lowest AIC among these nine loci. Results were based on 1 x 10^6 permutations using a multiple linear regression analysis.

Supplemental table 8. *SLCO1B1* SNP genotypes and their association with methotrexate clearance measured during the first course of methotrexate (during consolidation) as the phenotype.

		Discovery cohort (n=434)		Validation cohort (n=206)		Combined cohort (n=640)	
dbSNP ID	Alleles (A/B)	Coefficient[†]	P value	Coefficient[†]	P value	Coefficient[†]	P value
rs11045879	C/T	15.41	4.24×10^{-7}	5.20	0.117165	11.95	$3.19X\ 10^{-7}$
rs4149081	A/G	13.72	1.06×10^{-5}	5.31	0.110054	10.76	$5.63X\ 10^{-6}$
rs10841753	A/G	-8.17	0.005663	-9.95	0.001463	-8.22	0.000249
rs11045818	C/T	-9.34	0.006903	-14.28	8.90×10^{-5}	-10.10	0.000118
rs11045872	A/G	6.53	0.046869	8.43	0.01209	6.59	0.007549
rs2900476	C/T	8.13	0.003208	1.79	0.542274	5.79	0.005616
rs4149076	C/T	6.24	0.015978	-1.86	0.470871	3.60	0.062062
rs17328763	G/T	-5.65	0.089194	-7.72	0.019195	-5.72	0.020287
rs11045787	C/T	-5.41	0.100657	-7.24	0.028685	-5.36	0.029183
rs7966613	A/G	-6.42	0.012035	1.57	0.544463	-3.87	0.043761

[†] The coefficient or effect size represents the decrease (- value) in methotrexate clearance (ml/min/m^2) for those patients who carry an additional B allele. For instance, methotrexate clearance for patients in the discovery cohort who carry the C allele at the rs11045879 locus will have higher methotrexate clearance (15.41 ml/min/m^2 higher) than those patients that carry the T allele. For each additional C allele, methotrexate clearance will increase on average 15.41 ml/min/m^2.

Supplemental table 9. Association of the non-synonymous *SLCO1B1* SNP, T521C, with methotrexate clearance in the discovery, validation and combined patient cohorts.

			Discovery cohort (n=387)			*Validation cohort (n=102)*			*Combined cohort (n=489)*		
dbSNP ID	***Location***	***Alleles (A/B)***	***MAF***	***P value***	***Coefficient†***	***MAF***	***P value***	***Coefficient†***	***MAF***	***P value***	***Coefficient†***
rs4149056	non-synonymous (V134A)	T/C	0.12	1.9×10^{-7}	-14.2 (-19.1,-9.3)	0.13	0.40	-3.7 (-12.4, 3.4)	0.12	1.2×10^{-7}	-12.5 (-16.8, -8.4)

†The coefficient or effect size represents the decrease (- value) in methotrexate clearance (ml/min/m^2) for those patients who carry an additional C allele. For instance, methotrexate clearance for patients in the combined cohort who carry the *C* allele at the rs4149056 locus will have lower methotrexate clearance (-12.5 ml/min/m^2 lower) than those patients that carry the *T* allele. For each additional *C* allele, methotrexate clearance will decrease on average 12.5 ml/min/m^2. Confidence intervals are shown in parenthesis.

Supplemental table 10. Results from multivariate analyses in 489 patients with *SLCO1B1* T521C (rs4149056) and 500K SNPCHIP data. Covariates in the multivariate analyses included age, race, sex, methotrexate treatment regimen as well as genotype data for T521C and the top ranked *SLCO1B1* SNP rs11045879. Three separate models were conducted:1-2) either rs11045879 or T521C were included in the model individually (top two panels) and 3) both rs11045879 and rs4149056 were allowed to compete in the model simultaneously (bottom two panels). Shown are the coefficients and P values for T521C and rs11045879 SNPs.

Analysis	dbSNP ID	Alleles (A/B)	Coefficient[†]	P value
T521C and rs11045879 were included in the model individually	rs11045879	C/T	-10.9	4.9×10^{-11}
	T521C	T/C	-12.8	9.1×10^{-9}
T521C and rs11045879 were included in the model simultaneously	rs11045879	C/T	-9.8	0.006
	T521C	T/C	-3.6	0.36

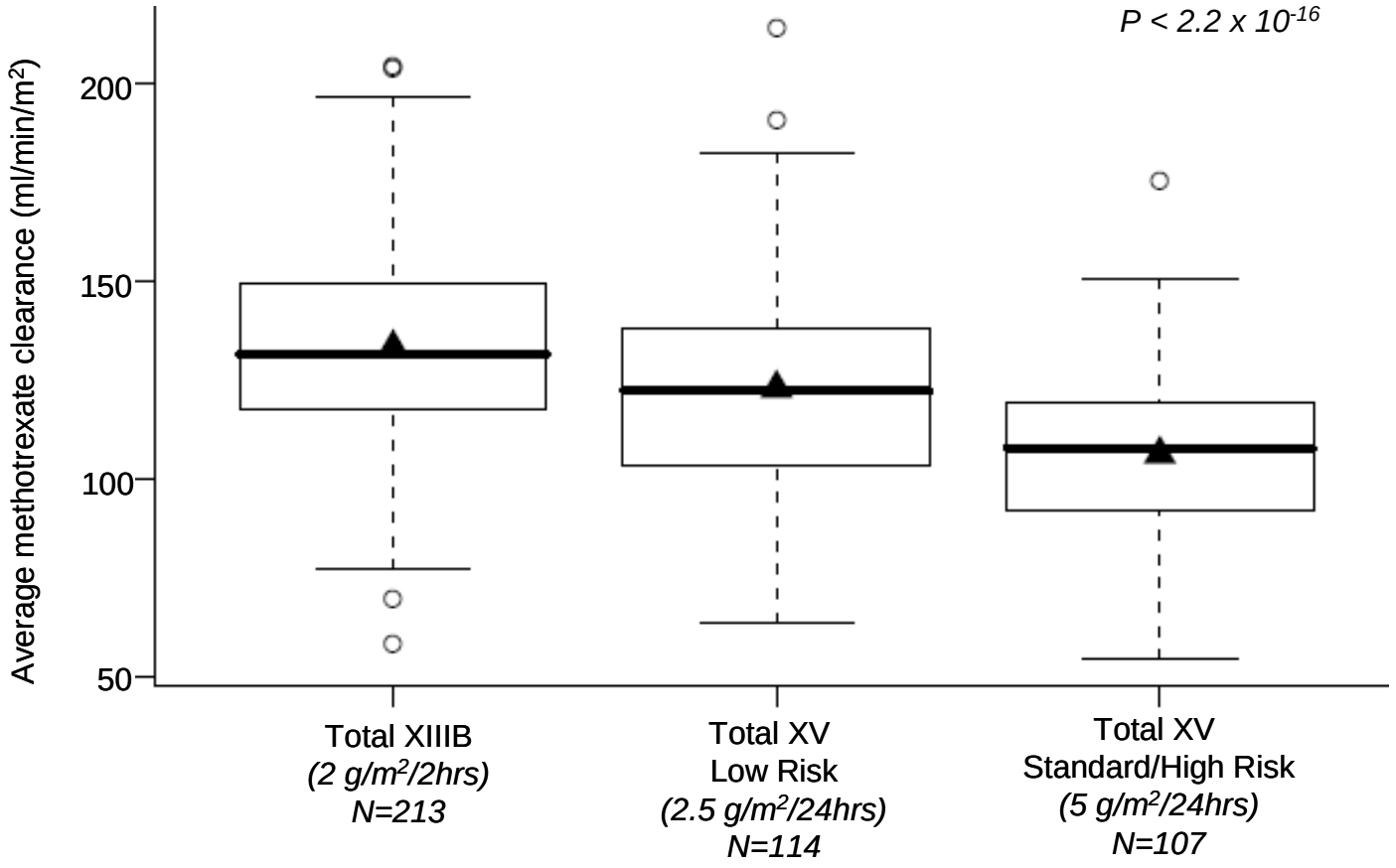
[†] The coefficient or effect size represents the increase (+ value) or decrease (- value) in methotrexate clearance (ml/min/m^2) for each variable listed. For example, patients with an additional copy of the T allele for rs11045879 will have a lower methotrexate clearance (-10.9 ml/min/m^2 lower) than those patients that carry a C allele. For each additional T allele, methotrexate clearance will decrease on average 10.9 ml/min/m^2

Supplemental table 11. Comparison of gene level analysis and single SNP analysis at the *SLCO1B1* locus. In the combined (discovery plus validation) gene level analysis (n=640), the top 10 components from principal component analysis using the genotypes of 27 SNPs annotated to the *SLCO1B1* gene explained 95.1% genotypic variation at the locus. If we rank the components by the percentage of variation explained, with the first principal component explaining the highest genetic variation, the third component is the component that most significantly correlates with methotrexate clearance ($P = 1.1 \times 10^{-13}$). The loadings (or weights) of each genotype for this component are listed below such that the loading with highest absolute value indicates the genotype at that SNP is contributing most to the component. No other components have a significant association with methotrexate clearance at the $P < 0.05$ level. A loading value of zero indicates that the particular polymorphism is not contributing to the principal component.

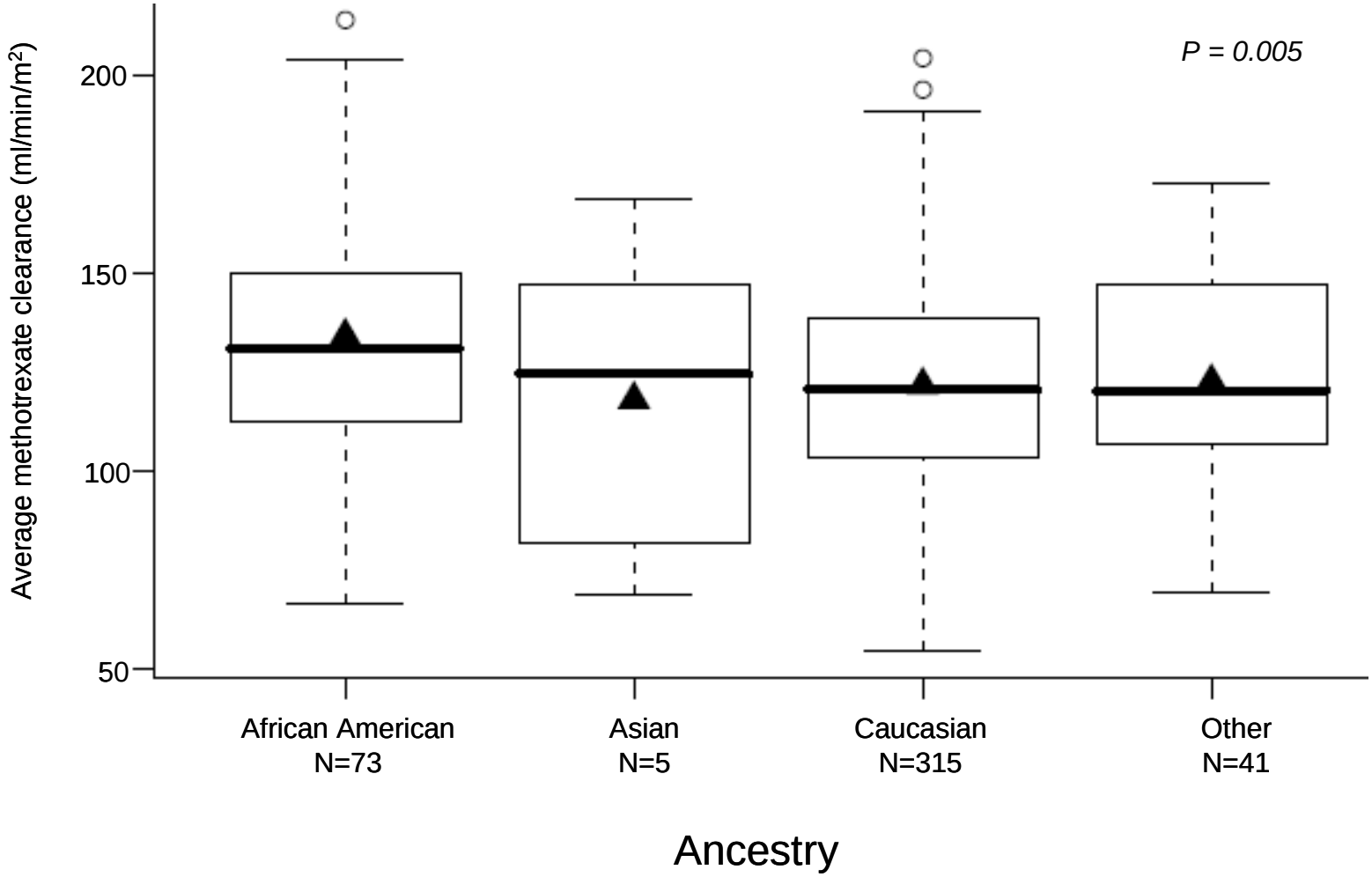
dbSNP ID	Loadings/weights for the significant principal component	Listed as significantly associated with methotrexate clearance in Table 1	Listed as significantly associated with methotrexate clearance in Table 2
rs11045879	-0.54	Yes	
rs4149081	-0.54	Yes	Yes
rs11045872	0.4	Yes	
rs11045818	-0.27	Yes	Yes
rs2900476	0.22	Yes	Yes
rs10841753	0.21	Yes	Yes
rs2291075	-0.15		
rs4149076	-0.13	Yes	
rs7139376	-0.13		
rs7966613	-0.12	Yes	
rs11045808	0.07		
rs991262	-0.05		
rs11045802	-0.03		
rs1463565	0.01		
rs999278	0.01		
rs2291076	-0.02		
rs4149069	-0.01		
rs852549	-0.01		

rs4149014	0		
rs2417955	0		
rs17328763	0	Yes	Yes
rs11045787	0	Yes	
rs7973095	0		
rs11045819	0		
rs11045776	0		
rs11045825	0		
rs7137959	0		

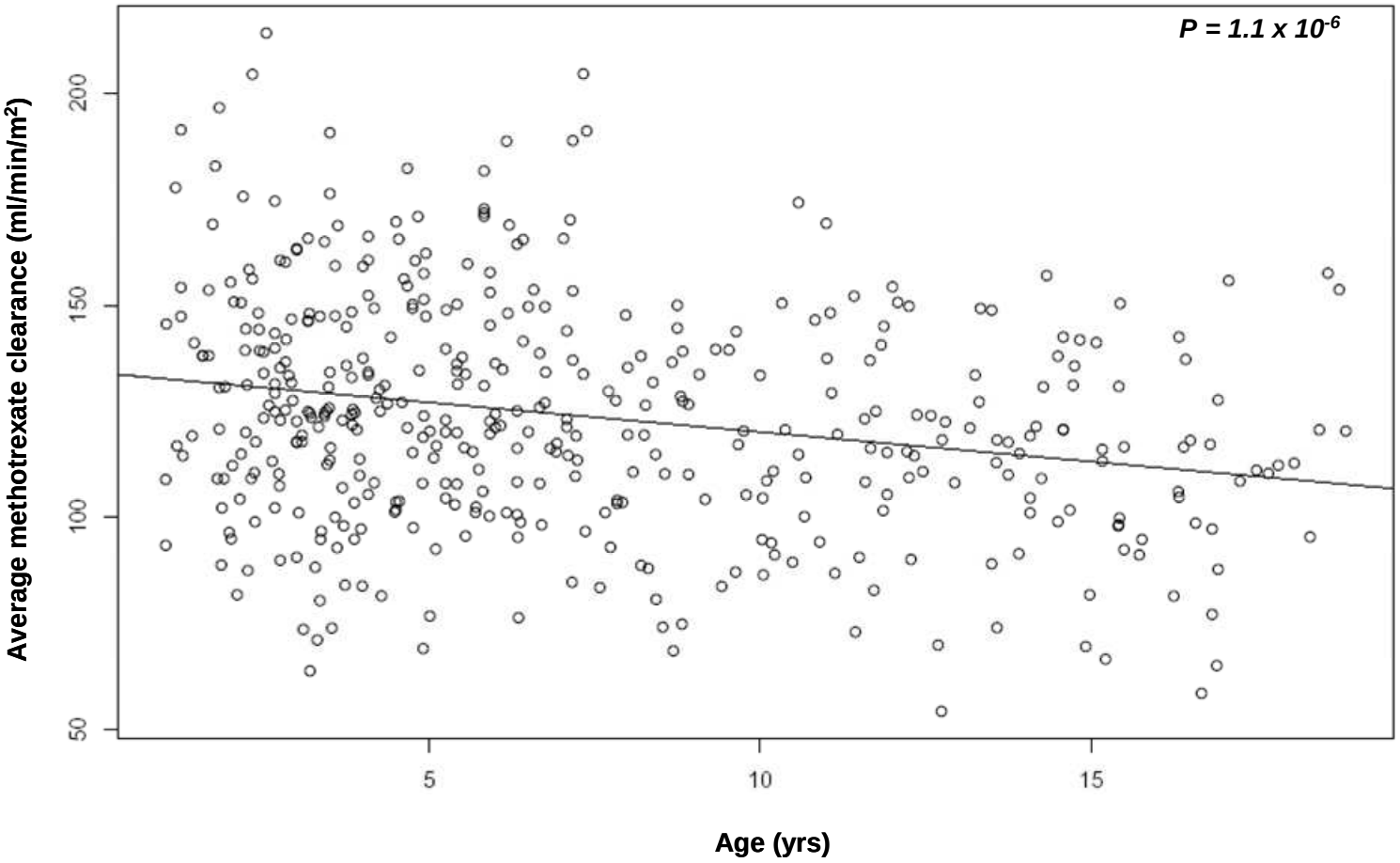
Supplemental figure 1. Average methotrexate clearance significantly differed among treatment regimens. Patients in the discovery cohort were treated on St. Jude Children's Research Hospital Total XIIIB (n= 213) and Total XV Low Risk (n=114) or Total XV Standard/High (n = 107) risk ALL protocols. Shown are quartiles, the minimum and maximum range of methotrexate clearance, outliers (open circles) and the mean value of methotrexate clearance (closed triangles). P values were generated using analysis of variance (ANOVA).



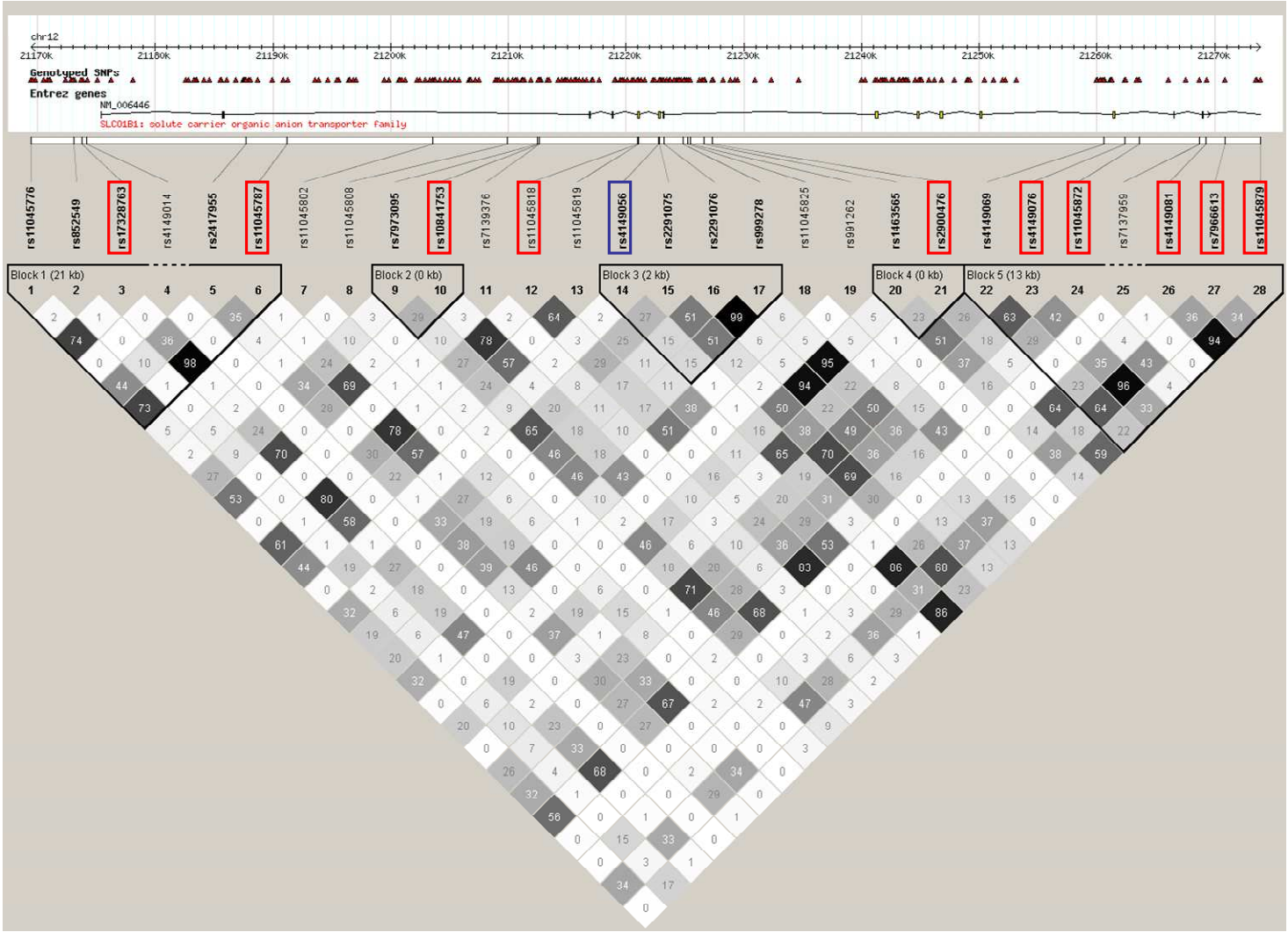
Supplemental figure 2. Average methotrexate clearance significantly differed among racial groups in the discovery cohort. Racial groups in the discovery cohort included Asian patients (n=5), African American patients (n=73), those designated as “Other” (n=41), and Caucasian patients (n=315). Patients were assigned to racial groups (Asian, African-American, and Caucasian) if they had >90% of the respective ancestry genotypes. Shown are quartiles, the minimum and maximum range of methotrexate clearance, outliers (open circles), and the mean value of methotrexate clearance (closed triangles). P values were generated using analysis of variance (ANOVA).



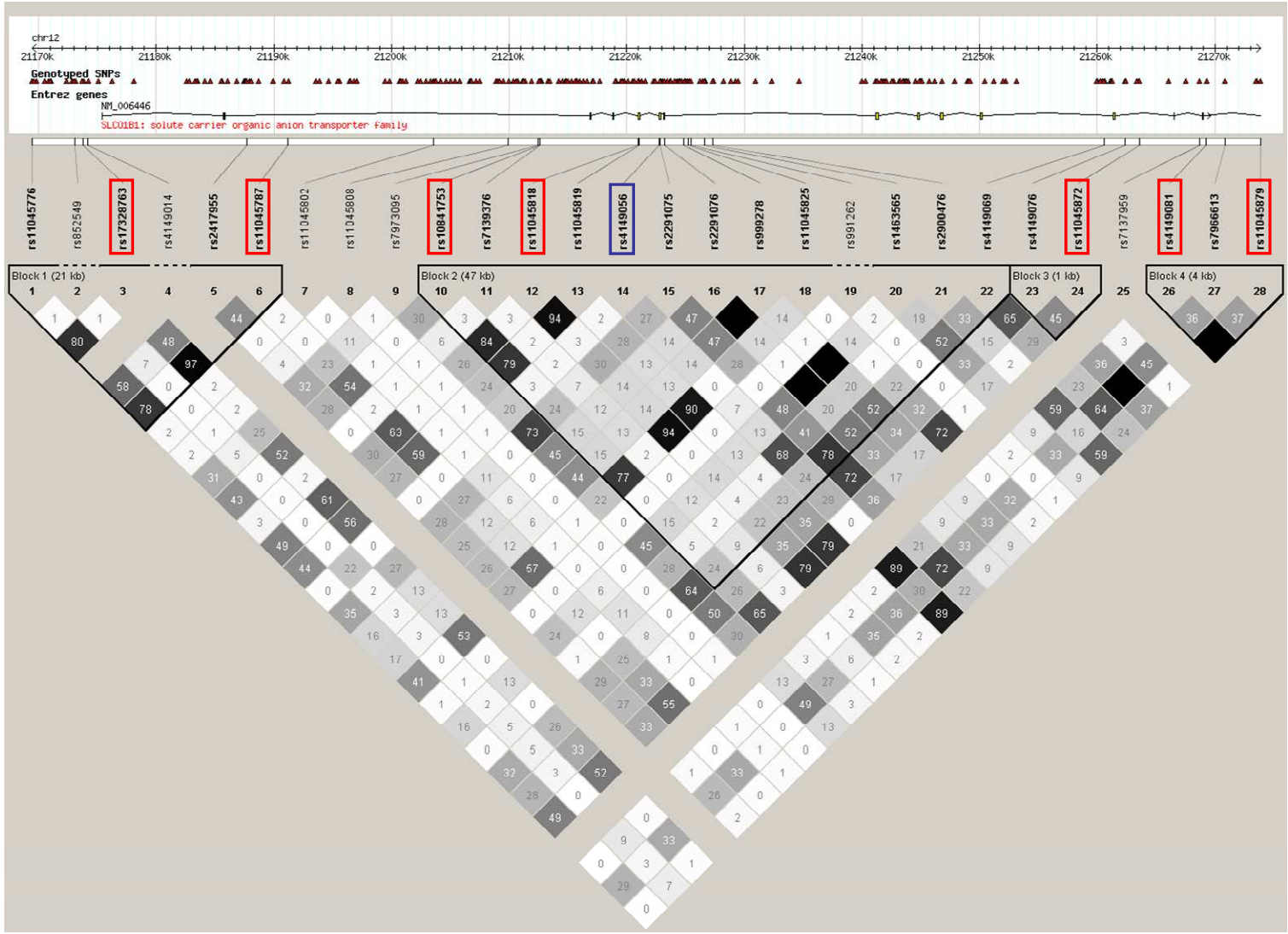
Supplemental figure 3. Average methotrexate clearance decreased with increasing age in the discovery cohort.



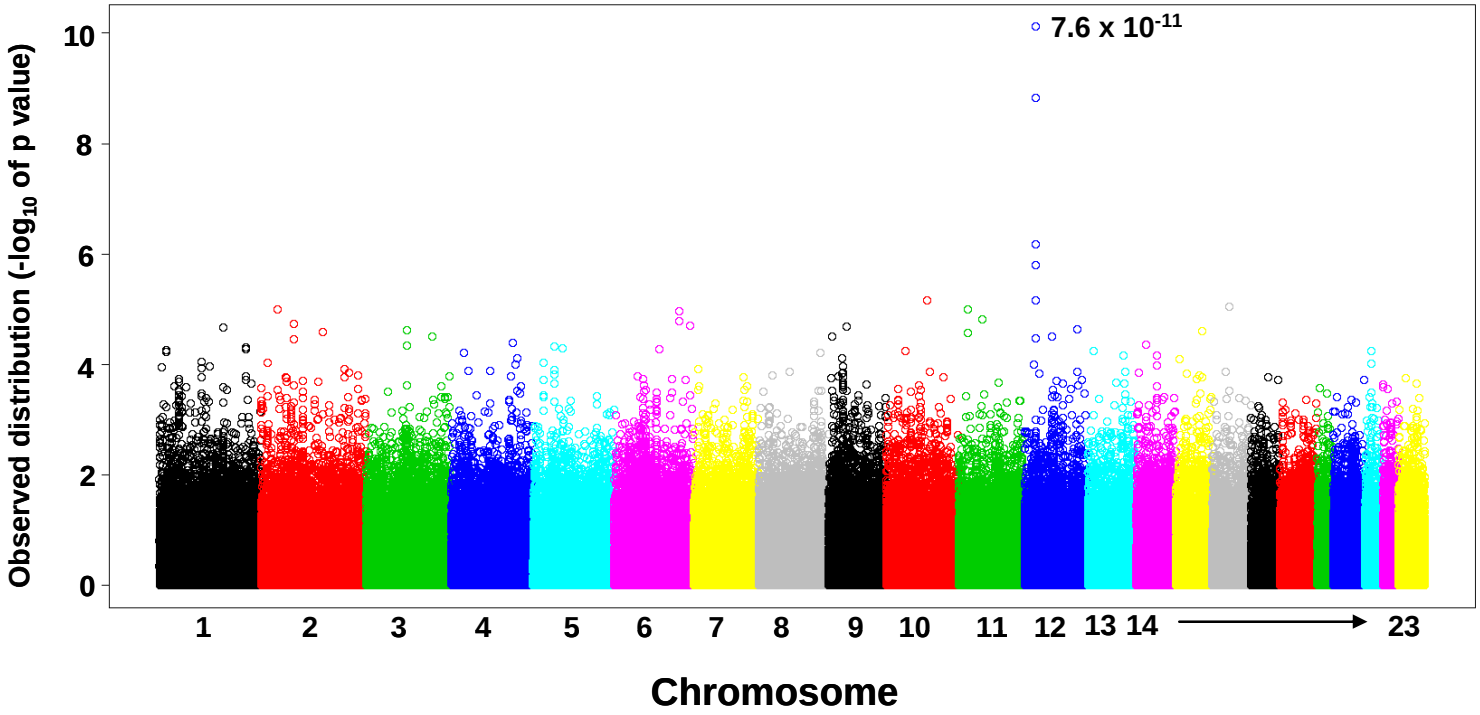
Supplemental figure 4A. Linkage disequilibrium (r squared) structure among *SLCO1B1* SNPs in the discovery cohort. Genotype data are from 285 patients of Caucasian ancestry from the study population. Shown are 27 *SLCO1B1* SNPs that were typed on the 500K Mapping Array sets plus rs4149056 (blue box, T521C). Boxed in red are the SNPs that were significantly (P < 0.05) associated with methotrexate clearance in the discovery cohort (n=434) (Table 1). There were 5 distinct linkage disequilibrium (LD) blocks within the region indicated that were generated using the Haploview software. Linkage disequilibrium (r squared) between SNPs increases from the color white (r^2=0; no evidence of LD), to shades of grey (0 < r^2 < 1), to the color black (r^2=1; strong evidence of LD).



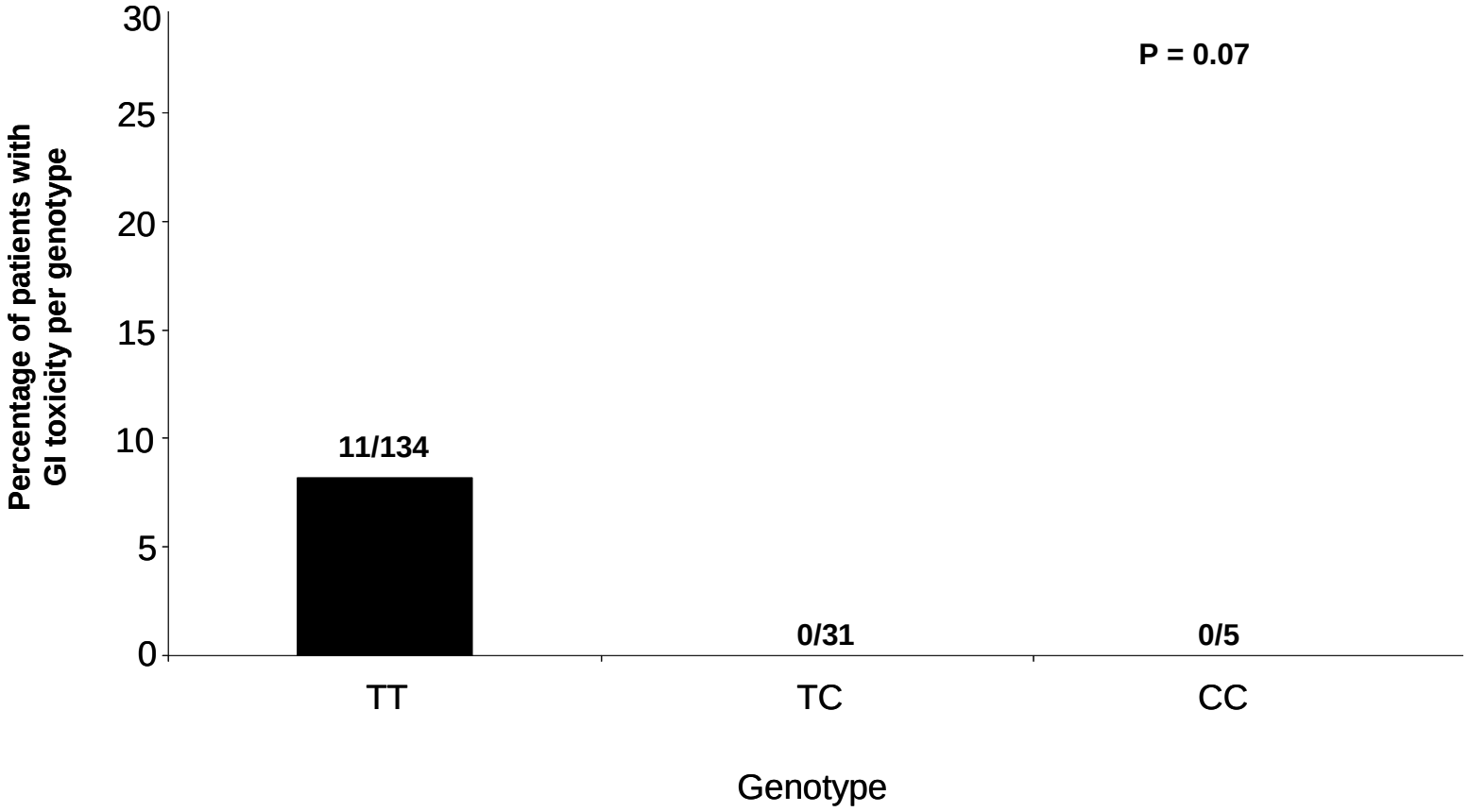
Supplemental figure 4B. Linkage disequilibrium (r squared) structure among *SLCO1B1* SNPs in the validation cohort. Genotype data are from 145 patients of Caucasian ancestry from the validation cohort. Shown are 27 *SLCO1B1* SNPs that were typed on the 500K Mapping Array set plus rs4149056 (blue box, T521C). Boxed in red are the SNPs that were significantly (P < 0.05) associated with methotrexate clearance in the validation cohort (n=206) (Table 1). There were 4 distinct linkage disequilibrium (LD) blocks within the region indicated that were generated using the Haploview software. Linkage disequilibrium (r squared) between SNPs increases from the color white (r^2=0; no evidence of LD), to shades of grey (0 < r^2 < 1), to the color black (r^2=1; strong evidence of LD).



Supplemental figure 5. Genome-wide P values showing the association of single nucleotide polymorphisms (SNPs) with methotrexate clearance in the combined (discovery plus validation) cohort of 640 children with acute lymphoblastic leukemia. Shown is the distribution of P values (as $-\log_{10}$ values) for the association of 398,699 SNP genotypes with methotrexate clearance in the combined cohort. Highlighted is the P value for the most significant association identified between a single SNP (located in *SLCO1B1* on chromosome 12) and methotrexate clearance.



Supplemental figure 6. Association between the non-synonymous *SLCO1B1* SNP, T521C, and the occurrence of gastrointestinal toxicity during the consolidation phase of treatment during St. Jude's Total XIIIB ALL protocol. The bar graph below displays the percentage of patients (plotted on the y-axis) per genotype (plotted on the x-axis) that suffered grade 3 to 4 gastrointestinal toxicity. Numbers above each bar represent the numbers of patients that did vs. did not suffer toxicity for the specified genotype.



References

1. Price,AL, Patterson, Plenge et al : Principal components analysis corrects for stratification in genome-wide association studies. Nat Genet. 38:904-909, 2006