Supplementary Methods

Blood sample analyses

Within 15 minutes of sample collection, each was centrifuged at 1200 rpm for 10 minutes at 4°C, transferred to polypropylene pour-off tubes, and then stored at -70°C. The melphalan concentration in each plasma sample was analyzed by liquid chromatography-tandem mass spectrometry. Briefly, plasma calibration samples (200 μL) containing various concentrations of melphalan (10 nM to 1 μM in standard solution) and/or hesperetin (100 nM) as internal standard were mixed with 1 mL acetonitrile for protein precipitation. After 1 minute of vortexing, the mixture was centrifuged at 11,000 x g for 10 minutes. Supernatants were transferred into clean tubes and evaporated to dryness under nitrogen stream. The residue was reconstituted with 100 μL of 5% methanol containing 0.1% formic acid. The reconstituted solution was centrifuged at 11,000 x g for 2 minutes, supernatant transferred to autosampler vials, and 20 μL injections were used for LC-MS/MS analysis. Mass transitions monitored were 305.01>245.94 for melphalan and 303.15>177.06 for hesperetin. The lower limit of quantification for melphalan was 10 nM (3.05 ng/mL equivalent). The method was validated per FDA criteria (1).

Population PK modeling

Melphalan pharmacokinetics was modeled using a nonlinear mixed-effects approach in NONMEM 7, version 7.1.2 (ICON Development Solutions; Ellicott City, Maryland). While PK samples were collected from 119 patients, one patient was excluded due to incorrect sampling record, and data from 118 patients were used to conduct population modeling. Melphalan concentration-time data were fitted with a two-compartment PK model. Population PK parameter estimates included CL, V_1, Q, and V_2, and the AUC was determined by CL and melphalan dose for individuals (AUC_{inf} = DOSE/CL) and C_{max} was predicted concentration at the end of infusion. IPV of PK parameters was evaluated using an exponential random effects model. Residual variability (ε) was described as a proportional residual random error model. To note, the additive residual term was negligible in our model.

The patients' characteristics that were collected and included as potential covariates were: age, body weight, height, body surface area, fat free mass, BMI, gender, serum creatinine, absolute neutrophil count, hemoglobin, hematocrit, platelets, white blood cells, blood urea nitrogen, bicarbonate, c-reactive protein, total bilirubin, and albumin (Supplementary Table 3). Fat free mass was calculated using total body weight (TBW) and BMI (2):

$$\text{Fat free mass (kg)} = \frac{9.27 \times 10^3 \times \text{TBW (kg)}}{6.68 \times 10^3 + 216 \times \text{BMI}} \quad \text{(male)}$$

$$\text{Fat free mass (kg)} = \frac{9.27 \times 10^3 \times \text{TBW (kg)}}{8.78 \times 10^3 + 244 \times \text{BMI}} \quad \text{(female)}$$

Creatinine clearance was calculated utilizing TBW and the Cockcroft & Gault equation (3), CKD-EPI equation (4), or 24-hour urine data. Single covariates were screened to determine their contributions to inter-patient variability. Those individual covariates having significant influence ($p < 0.05$) on melphalan PK were then added stepwise to the structural model until no further significant reduction in OFV was detected. Following this, backward stepwise elimination from the full PK model was performed. The model was evaluated by comparing the difference in the OFV (ΔOFV), goodness of fit plots, and standard error of parameter estimates. All PK results were summarized using appropriate descriptive statistics.

Model accuracy and stability were evaluated via bootstrap re-sampling. For each of 1000 bootstrap iterations, model parameters were estimated and 95% confidence intervals of the bootstrap replicates were compared with parameter estimates from the final PK model. Model based simulation was then performed to evaluate the predictive performance of the final model using VPC from 1,000 simulations.

*PK studies statistical analyses***:** First-order conditional estimation (FOCE) was used to estimate parameters. Development of the PK structural model and the stepwise selection of covariates were performed based on the changes in the OFV, such that a decrease in OFV greater than 3.84 was indicative of significance ($p < 0.05$, degree of freedom = 1) using the log likelihood ratio test. Graphic analyses were performed using R, version 3.1.1 (5).

Ex vivo sensitivity to melphalan

Cytotoxicity levels (IC_{50} values) were measured by cell proliferation assay using a standard water-soluble tetrazolium salt (WST-1) reagent (Roche Applied Science, Germany). Briefly, PBMCs were seeded in 96-well plates and exposed to various concentrations of melphalan (0, 1, 3, 5, 10, 30, 50, 75, and 150 μM). After incubation at 37°C for 24 hours, 10 μL of the ready-to-use WST-1 reagent was added. The absorbance at 450 nm was then measured, and IC_{50} values were determined using Sigma Plot 12.0 software (Systat Software Inc). Experiments were conducted in triplicate and the mean IC_{50} values were used for further statistical analyses.

SLC7A5 genotyping

Genomic DNA from untreated PBMCs was isolated using AllPrep DNA/RNA Mini Kit (cat. #80204, QIAGEN). The *SLC7A5* polymorphism, rs4240803, was determined using a TaqMan pre-validated genotyping assay (Catalog # C__1228783_10, Thermo Fisher Scientific, Waltham, MA, USA).

References

(1) FDA. Guidance for Industry: Bioanalytical Method Validation. Draft Guidance. (US Department of Health and Human Services, Food and Drug Administration, CDER and CVM, 2013).

(2) Janmahasatian, S., Duffull, S.B., Ash, S., Ward, L.C., Byrne, N.M. & Green, B. Quantification of lean bodyweight. *Clin. Pharmacokinet.* **44**, 1051-1065 (2005).

(3) Cockcroft, D.W. & Gault, M.H. Prediction of creatinine clearance from serum creatinine. *Nephron* **16**, 31-41 (1976).

(4) Levey, A.S. *et al.* A new equation to estimate glomerular filtration rate. *Ann. Intern. Med.* **150**, 604-612 (2009).

(5) R Core Team. *R: A language and environment for statistical computing*. (R Foundation for Statistical Computing, Vienna, Austria, 2016).

Supplementary Table 1. Stepwise selection of covariates in melphalan population pharmacokinetic model: A) forward selection and B) backward elimination

A)

Model	∆OFV	p value
1. Base model	0	-
2. 1 + Creatinine clearance on CL	-51.043	<0.0001
3. 2 + SLC7A5 genotype on V2	-15.123	<0.0001
4. 3 + Fat free mass in allometric scale	-27.665	<0.0001
5. 4 + Hematocrit on CL	-4.808	<0.05

B)

Model	∆OFV	p value
Full model	0	-
- (Creatinine clearance on CL)	41.826	<0.0001
- (SLC7A5 genotype on V2)	10.958	<0.01
- (Fat free mass in allometric scale)	28.039	<0.0001
- (Hematocrit on CL)	4.808	<0.05

Supplementary Table 2. World Health Organization Criteria for Oral Mucositis

GRADE	DESCRIPTION
0 (none)	None
I (mild)	Oral soreness, erythema
II (moderate)	Oral erythema, ulcers, solid diet tolerated
III (severe)	Oral ulcers, liquid diet only
IV (life-threatening)	Oral alimentation impossible

Supplementary Table 3. Patient clinical characteristics in pharmacokinetic study (male=69, female=50)

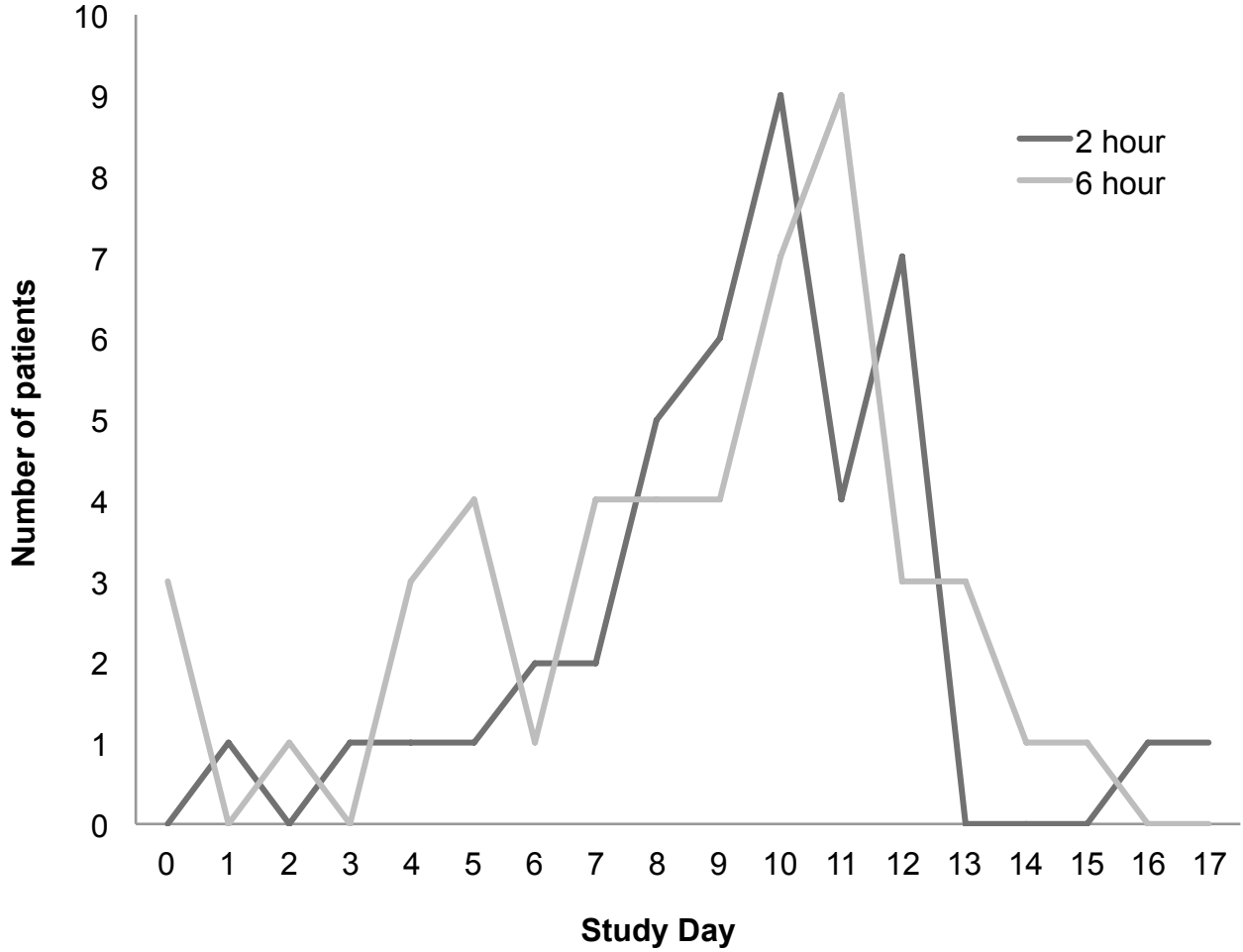
Characteristic	**Median**	**Range**
Age (years)	59	35 - 72
Weight (kg)	84.05	45.41 - 145.33
Height (m)	1.7	1.45 - 1.92
Body surface area (m^2)	1.96	1.37 - 2.50
Serum creatinine (mg/dL)	0.84	0.34 - 14.50
Fat free mass (kg)	59.9	31.3 - 81.9
Body mass index (kg/m^2)	29.24	17.34 - 48.73
Absolute neutrophil count (K/μL)	3.5	0.7 - 13.8
Hematocrit (%)	32.5	20.6 - 44.6
Hemoglobin (g/dL)	10.9	7.0 - 14.7
Platelet (L/μL)	188	41 - 420
White blood cell count (K/μL)	4.9	1.7 - 18.3
Blood urea nitrogen (mg/L)	14	5 - 59
Bicarbonate (mmol/L)	27	21 - 33
Blood glucose – non fasting (mg/dL)	96	71 - 233
C-reactive protein (mg/L)	2.97	0 - 56.3
Albumin (g/dL)	3.5	2.2 - 4.7
Total protein (g/dL)	6.1	4.3 - 10.8
Bilirubin (total) (mg/dL)	0.5	0.2 - 1.5

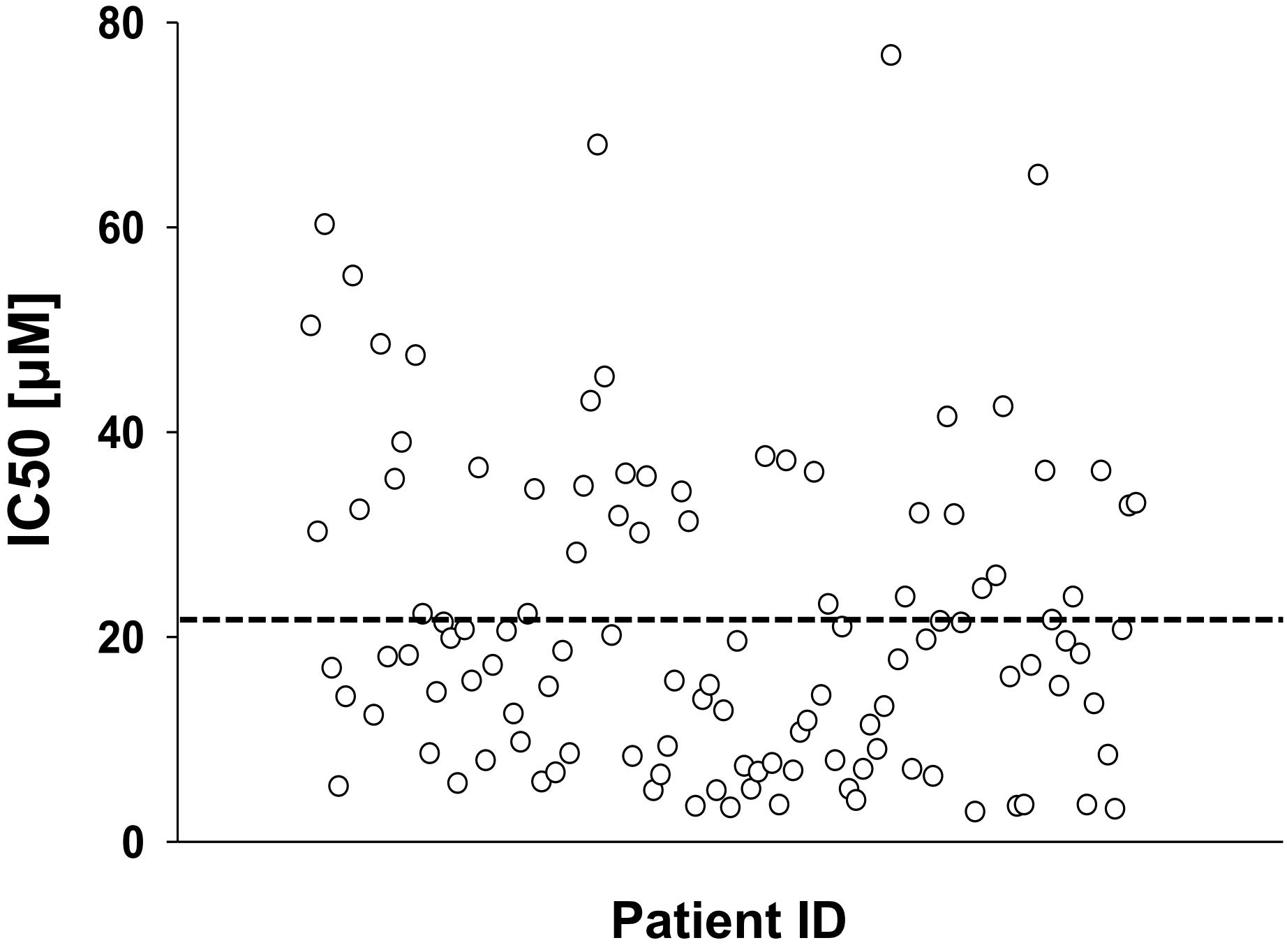
SUPPLEMENTARY FIGURE LEGEND

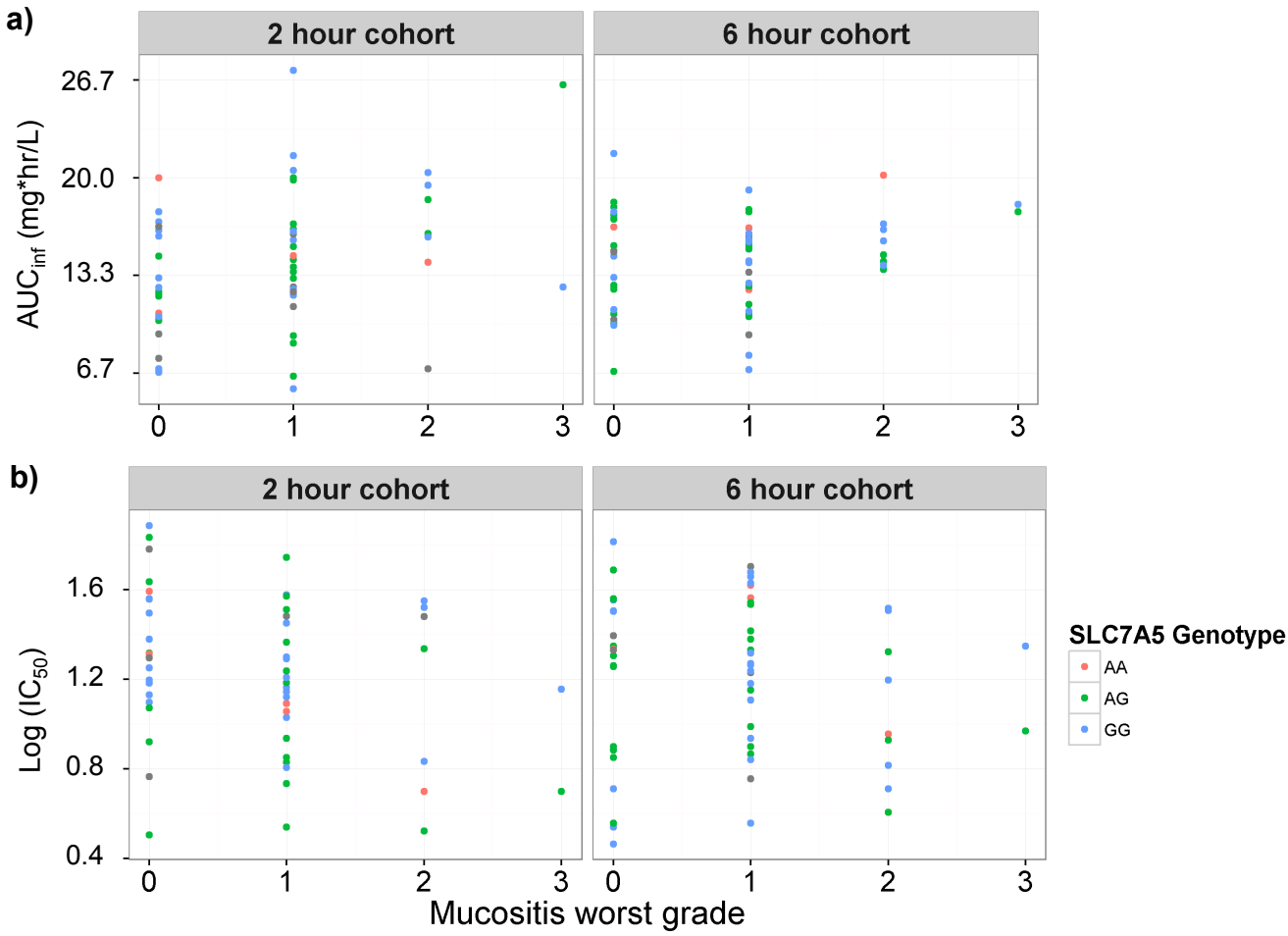
Supplementary Figure 1. Days to worst grade oral mucositis by cohort.

Supplementary Figure 2. Scatter plot of IC_{50} values of melphalan in PBMCs (N = 115). The dashed line represents the population mean value (21.4) of all IC50s.

Supplementary Figure 3. Individual plots of a) mucositis versus melphalan AUC values (N = 118) and b) mucositis versus *ex vivo* IC_{50} of PBMCs (N=115). The color indicates *SLC7A5* genotypes. Gray dots represent missing genotypes.


10
9
8
7
6
5
4
3
2
1
0
Number of patients
2 hour
6 hour
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17
Study Day


80
60
40
20
0
IC50 [µM]
Patient ID


a)
2 hour cohort
6 hour cohort
AUC$_{inf}$ (mg*hr/L)
26.7
20.0
13.3
6.7
0
1
2
3
b)
2 hour cohort
6 hour cohort
Log (IC$_{50}$)
1.6
1.2
0.8
0.4
Mucositis worst grade
SLC7A5 Genotype
AA
AG
GG