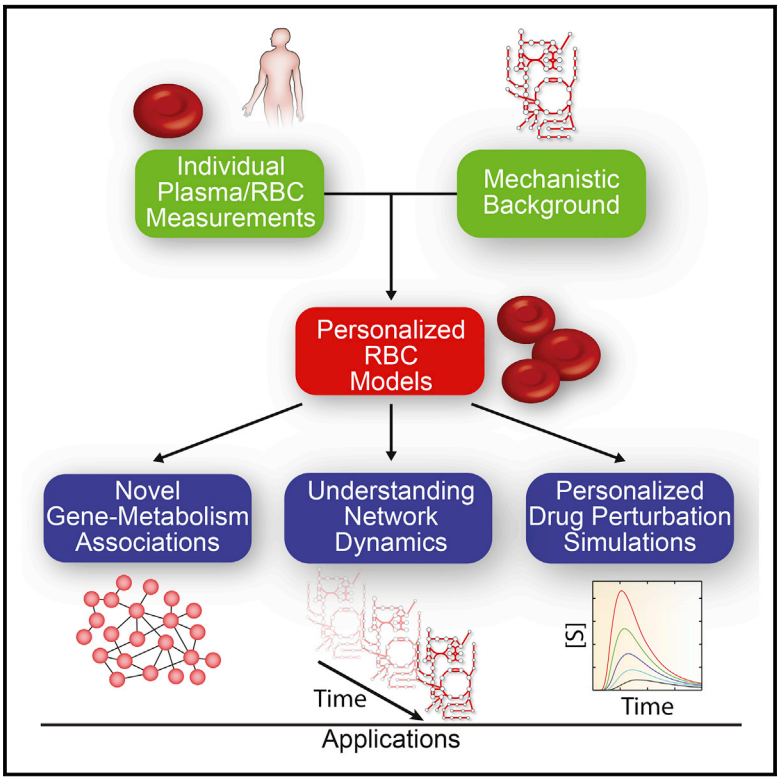
Article

Cell Systems

Personalized Whole-Cell Kinetic Models of Metabolism for Discovery in Genomics and Pharmacodynamics

Graphical Abstract



Authors

Aarash Bordbar, Douglas McCloskey, Daniel C. Zielinski, Nikolaus Sonnenschein, Neema Jamshidi, Bernhard O. Palsson

Correspondence

neema@ucsd.edu

In Brief

Bordbar et al. integrate plasma and erythrocyte metabolite measurements of 24 individuals with a metabolic network to construct personalized kinetic models. Estimated kinetic parameters better represent the individual's genotype than metabolite levels. Subsequently, the models are used to predict drug side effect susceptibility.

Highlights

- Personalized kinetic models of erythrocytes are built using metabolomic data
- Calculated kinetic parameters better represent the genotype than metabolite levels
- Individual differences in dynamics occur on physiologically relevant timescales
- Model simulations identify a potential mechanism for ribavirin-induced anemia

Bordbar et al., 2015, Cell Systems *1*, 283–292
October 28, 2015
http://dx.doi.org/10.1016/j.cels.2015.10.003

CellPress

Cell Systems
Article

Personalized Whole-Cell Kinetic Models of Metabolism for Discovery in Genomics and Pharmacodynamics

Aarash Bordbar,[1,6] Douglas McCloskey,[1] Daniel C. Zielinski,[1] Nikolaus Sonnenschein,[2] Neema Jamshidi,[1,3,5,*] and Bernhard O. Palsson[1,2,4]
[1]Department of Bioengineering, University of California, San Diego, La Jolla, CA 92093-0412, USA
[2]The Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, 2970 Hørsholm, Denmark
[3]Institute of Engineering in Medicine, University of California, San Diego, La Jolla, CA 92093-0412, USA
[4]Department of Pediatrics, University of California, San Diego, La Jolla, CA 92123-5109, USA
[5]Department of Radiological Sciences, University of California, Los Angeles, Los Angeles, CA 90095-1721, USA
[6]Present address: Sinopia Biosciences Inc., 600 West Broadway, Suite 700, San Diego, CA 92101, USA
*Correspondence: neema@ucsd.edu
http://dx.doi.org/10.1016/j.cels.2015.10.003

SUMMARY

Understanding individual variation is fundamental to personalized medicine. Yet interpreting complex phenotype data, such as multi-compartment metabolomic profiles, in the context of genotype data for an individual is complicated by interactions within and between cells and remains an unresolved challenge. Here, we constructed multi-omic, data-driven, personalized whole-cell kinetic models of erythrocyte metabolism for 24 healthy individuals based on fasting-state plasma and erythrocyte metabolomics and whole-genome genotyping. We show that personalized kinetic rate constants, rather than metabolite levels, better represent the genotype. Additionally, changes in erythrocyte dynamics between individuals occur on timescales of circulation, suggesting detected differences play a role in physiology. Finally, we use the models to identify individuals at risk for a drug side effect (ribavirin-induced anemia) and how genetic variation (inosine triphosphatase deficiency) may protect against this side effect. This study demonstrates the feasibility of personalized kinetic models, and we anticipate their use will accelerate discoveries in characterizing individual metabolic variation.

INTRODUCTION

Biofluid metabolite profiling has been an effective means for screening and diagnosis of numerous medical diseases (Lanpher et al., 2006; Pollitt et al., 1997). It is anticipated that more in-depth metabolomic profiling will provide even greater insight into diseases (Holmes et al., 2008; Psychogios et al., 2011). High-throughput technologies for measuring biomolecules (Beckonert et al., 2007; Tautenhahn et al., 2012) have aided in characterizing human variation and driving personalized diagnostics and therapeutics (Wishart et al., 2013). The decreasing costs of high-throughput data generation are enabling the ability to shift focus from population-based studies to individuals, and they even allowed researchers to track a single individual over a 14-month period (Chen et al., 2012).

Such advances are notable but often do not address interactions occurring within a complex metabolic network of a cell or its multi-compartment interactions with the blood stream. Physiology and physiological responses are multi-factorial, so, for the majority of diseases, a single data type will not be sufficient for diagnosing, prognosing, and determining appropriate treatment. There is a strong need for a systematic methodology to integrate complex datasets in a biologically coherent fashion for understanding measured metabolic and genetic variations in mechanistic terms.

Genome-scale constraint-based models (Orth et al., 2010) are a framework for multi-omic data integration that provides a deeper understanding of metabolic functions than data alone (Bordbar et al., 2014). Initial studies using these models to interpret individual variation are encouraging, addressing loss-of-function mutations, such as those found in Mendelian inherited disorders, or the complete absence of proteins (Agren et al., 2014; Jamshidi et al., 2011; Jamshidi and Palsson, 2006; Shlomi et al., 2009), as well as the integration of gene expression data for predicting selective cancer drug targets (Yizhak et al., 2014). However, the majority of human metabolic variation is subtler, affecting metabolite and enzyme levels as well as enzyme activity. These changes can affect cellular dynamics, which cannot be captured by constraint-based models. Whole-cell kinetic models can be developed from constraint-based models and can explicitly describe metabolite and enzyme properties (Jamshidi and Palsson, 2008a; Stanford et al., 2013). Previous classical kinetic models have been useful in interpreting individual variation in a mechanistic fashion (Jamshidi et al., 2002; Saucerman et al., 2004), demonstrating the potential to complement or exceed strictly statistic-based approaches. However, traditional kinetic models require extensive in vitro parametrization and, thus, are often impractical for studying large cohorts. A data-driven approach is needed to study personalized cellular dynamics.

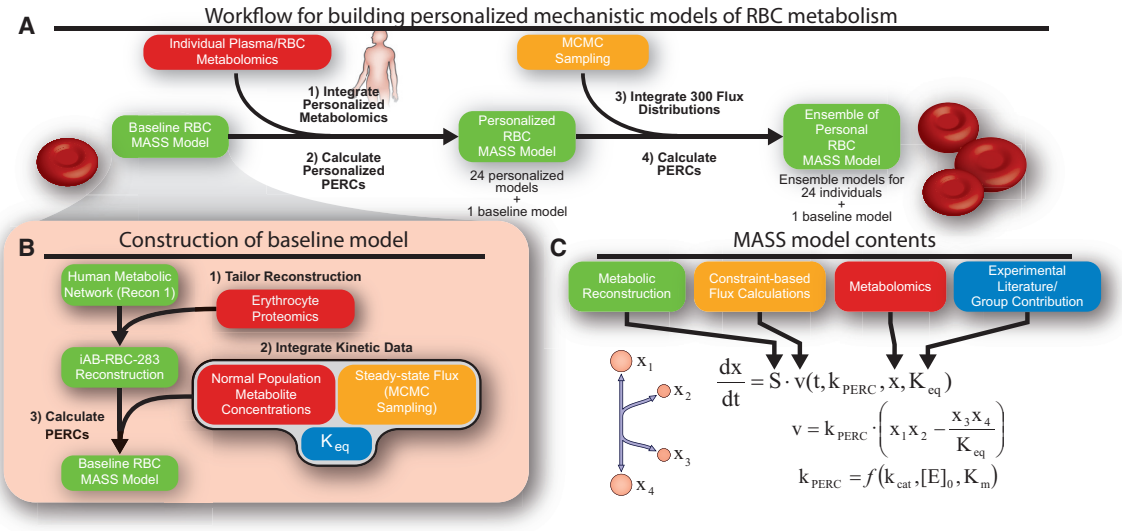


Figure 1. The MASS Approach for Building High-Throughput Personalized Kinetic Models

(A) A baseline human erythrocyte model was tailored to 24 individuals' plasma and erythrocyte metabolomics. To address inter-individual flux variability, 300 models for each individual were built with candidate flux states from MCMC sampling.

(B) To build the baseline kinetic model, we tailored a global human metabolic network for erythrocyte metabolism using published proteomic data. The network was integrated with a baseline flux state (from MCMC sampling), K_{eq}s (from literature, when available, and computational estimations), and normal population metabolite levels (from literature).

(C) The contents and equations of a MASS model are shown.

In this study, using plasma and intracellular erythrocyte metabolomics, we constructed a collection of personalized whole-cell kinetic models of erythrocyte metabolism for 24 healthy individuals. The method is a scalable means to integrate individual metabolite measurements with the underlying metabolic network to determine enzyme activity, providing a more direct means to compare the genotype with protein function. Analysis of the models indicates that enzymatic kinetic parameter variation reflects genetic variation. Further, a functional kinetic model was constructed, enabling simulation of perturbations for direct physiological interpretation of the consequences of detected metabolic variation in response to drug exposure.

RESULTS

Network-Based Kinetic Description of Omic Data

Traditionally, kinetic models of biochemical reaction networks are constructed through a biophysical approach using rate laws that require in-vitro-derived kinetic parameters (such as Michaelis constant, K_m; turnover number, k_{cat}; and enzyme concentration, E_0). The difficulty in obtaining kinetic parameters and their appropriateness for in vivo situations (Teusink et al., 2000) limits scaling of traditional kinetic models to large networks, let alone to a large number of individuals. In recent years, there have been numerous approaches for building kinetic models for metabolic, signaling, and gene regulatory networks (Chakrabarti et al., 2013; Danino et al., 2010; Khodayari et al., 2014; O'Dea et al., 2007; Stanford et al., 2013). In particular, one approach applies the law of mass action, which is a fundamental principle in chemistry and ultimately is the basis and derivation of mechanistic reaction descriptions. In this study, we applied the Mass Action Stoichiometric Simulation (MASS) approach (Jamshidi and Palsson, 2010), which allows for flexibility in incorporating mechanisms in varying degrees of detail, to construct personalized models using metabolomic profiles (Figure 1C).

Briefly, a MASS model is comprised of the network topology, derived from a metabolic network reconstruction (Feist et al., 2009; Oberhardt et al., 2009), and high-throughput data representing the steady state of the network (Figure 1C). Enzyme kinetics of the metabolic reactions are approximated by a phenomenological kinetic constant, the pseudo-elementary rate constant (PERC). The PERC represents a combination of traditional kinetic parameters and is consistent with all condition-specific omic data used. A mathematical comparison of MASS and rate law kinetic models is presented in the Supplemental Experimental Procedures, including simulations showing that individual variability is the overriding factor of dynamic simulations and not the choice of modeling approach. Thus, MASS models are a practical approximation of network dynamics based on an observable homeostatic state. Workflows, advantages, disadvantages, and examples of MASS models have been discussed previously (Jamshidi and Palsson, 2008a; Jamshidi and Palsson, 2010; Palsson, 2011) and demonstrated for cellular models of metabolism. This approach has not yet been implemented for the simultaneous construction of a library of individual-specific personalized models using in vivo omic data.

Workflow for Personalized Kinetic Model Construction

The workflow to build personalized erythrocyte kinetic models is shown in Figures 1A and 1B. We first built a baseline model using the average metabolite concentrations of the human plasma and erythrocyte (Figure 1B). A metabolic network reconstruction of the human erythrocyte (iAB-RBC-283, Bordbar et al., 2011) was derived from the global human metabolic network (Duarte

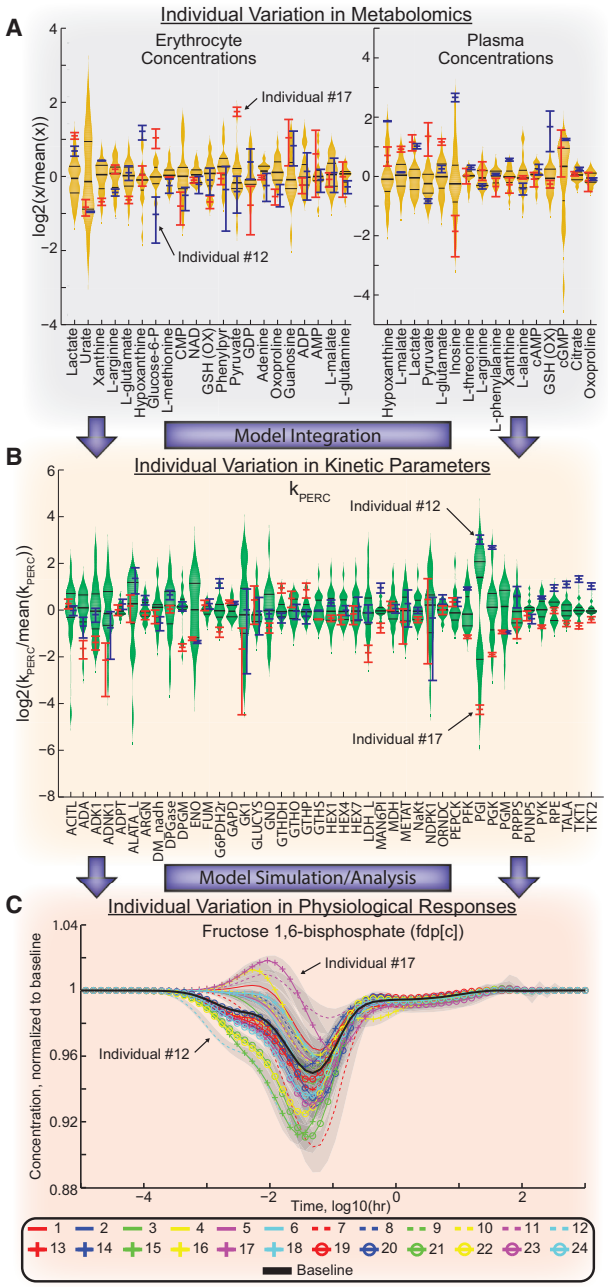


Figure 2. Personalized MASS Models Integrate Omic Data to Obtain System-Level Understanding of Individual Variation

Two individuals (12 and 17) with differing simulated responses to a redox load are highlighted with confidence intervals (mean ± SD).

(A) Violin plots of significantly different erythrocyte and plasma metabolite concentrations accounted for in the model ($p < 0.05$, Bonferroni corrected, ANOVA). 39 of the 69 erythrocyte and 39 of the 44 plasma metabolites were significantly different between at least two individuals, though not necessarily individuals 12 and 17. Black lines on the violin plots represent the 25th, 50th, and 75th quantiles.

(B) Metabolite levels are integrated to obtain network kinetic parameters. Violin plots of significantly different PERCs of intracellular reactions are presented. 54 PERCs in the model were significantly different between at least two individuals.

(C) An illustrative example of how personalized kinetic models can be simulated to assess systemic differences to perturbations. 25 metabolites and 36 reaction fluxes had significantly different responses between individuals when the initial NADPH/NADP ratio was perturbed. For example, fructose 1,6-bisphosphate levels responded quite differently between individuals. See the Supplemental Experimental Procedures and Figure S4 for simulation details.

et al., 2007) with the compilation and integration of multiple proteomic datasets obtained from the human erythrocyte (Figure 1B). The iAB-RBC-283 was then tailored to the scope of the metabolomic data obtained in this study (Data S1). Average metabolic concentrations (x) and equilibrium constants (K_{eq}) were determined from experimental literature. Equilibrium constants not available in the literature were determined through computational methods (Noor et al., 2012). Reaction fluxes (v) were determined by Markov chain Monte Carlo (MCMC) sampling (Bordbar et al., 2010; Schellenberger et al., 2011) of the constraint-based model, where the lower and upper bounds of exchange fluxes were parameterized by available experimental data (Data S1). After the various data (v, x, and K_{eq}) were integrated, PERCs (k_{PERC}) could be calculated for all reactions by solving the equation for the one missing value (Figure 1C). The flux state was estimated based on available physiological data through constraint-based modeling and was independent of the PERC estimates, avoiding potential concerns of overly conservative estimates from mass action constraints (Chakrabarti et al., 2013). The final model and its characteristics are presented in Figure S1.

Next we measured fasting-state plasma and erythrocyte metabolites for 24 healthy individuals (Figure 1A). Further, measurements were made at four different time points in four individuals over a 6-week period, allowing us to compare the same individual across multiple measurements to serve as controls. Overall, we measured 44 plasma and 69 intracellular erythrocyte metabolites, with 76 of the 113 metabolites being present in the MASS model. We first assessed whether the plasma and intracellular measurements were redundant. For example, is measuring just the plasma concentration of a particular metabolite enough to know the concentration of the intracellular erythrocyte metabolite? We confirmed that the plasma and intracellular metabolite measurements were not redundant, and measuring intracellular levels is important for assessing the metabolic state of even a blood cell. Of the 39 metabolites that were measured in both the plasma and the erythrocyte, only five (lactic acid, hypoxanthine, L-threonine, L-tryptophan, and uridine) were correlated between the two compartments ($R^2 > 0.5$, $p < 0.05$, Pearson correlation).

We constructed a personalized kinetic model for each individual by substituting the average of the measured metabolite concentrations of the individual under consideration for the baseline levels and calculating a personalized set of PERCs by re-solving all the equations. For the baseline model, a single flux reaction state was used. However, individuals might have different baseline reaction flux states. To address inter-individual flux variability, we utilized an ensemble approach (Tan and Liao, 2012) where 300 candidate flux distributions from the MCMC sampling were used to build 300 models for each individual. Introducing uncertainty into the reaction flux states mitigated the lack of direct flux measurements and allowed us to determine significantly different distributions of PERCs between ensembles of kinetic models for individuals, regardless of flux state (Figure 2B).

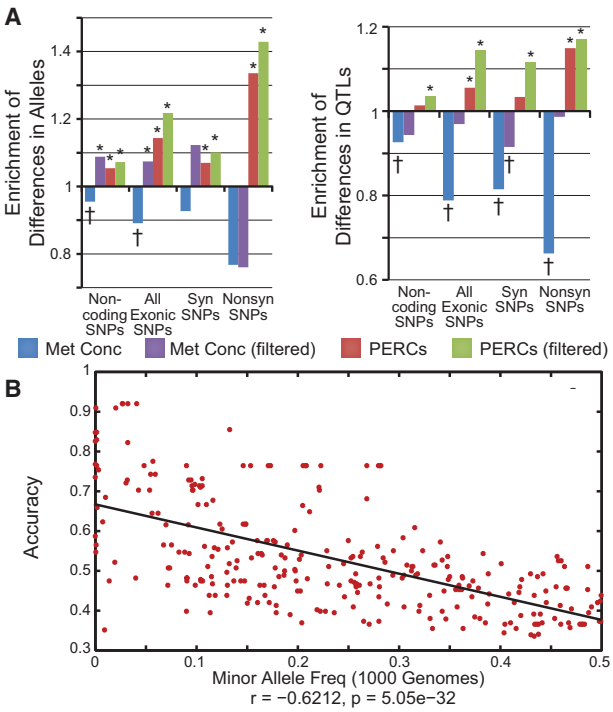


Figure 3. Personalized Kinetic Models Can Be Used for Genomic Discoveries

(A) Significantly different PERCs between individuals are enriched with allelic differences in the region of the associated genes (left). Removing spurious PERCs determined from controls increased levels of enrichment (filtered PERCs). Significantly different metabolite levels are often depleted with allelic differences in the gene regions whose associated reactions are connected to the metabolite. Allelic differences also are enriched with significantly different PERCs (right) and are significantly depleted with significantly different metabolites that are directly linked to reactions associated with the genes. Comparing metabolite differences to only the most representative reaction (filtered metabolite concentrations) connected to that metabolite did artificially boost results, as expected. However, the results were still not as representative of the genotypes as those found with the PERCs (*significantly enriched, †significantly depleted; $p < 0.05$, binomial distribution).

(B) Accuracy of genotype variations as predictors of significantly different PERCs was significantly correlated (Pearson correlation) with the minor allele frequency of the SNPs.

Further, we assessed the robustness of PERC calculations in terms of technical variability in metabolomic measurements (Supplemental Experimental Procedures; Figure S2).

Model Parameters Are Associated with the Genotype

The PERC is a bulk approximation of traditional enzymatic parameters (K_m, k_{cat}, and E_0) (Figure 1C). Individual K_m and k_{cat} depend on the protein sequence and the protein-coding exon DNA sequence. Variation in E_0 depends on other non-coding sequence variations (i.e., variants in promoters, enhancers, synonymous mutations in exons, and splice junctions). Thus, calculated PERCs also should depend on genetic variants. To assess this dependence, we focused on genetic variants occurring in the gene region of the metabolic genes (exon, intron, and promoter regions) in the constructed erythrocyte kinetic model, thus ignoring variants outside the gene region that may potentially affect E_0. The variants in the chosen region have a higher chance of being functionally relevant to enzyme kinetic differences.

To test the biological relevance of PERCs, we determined if ensembles of PERCs that were significantly different (Experimental Procedures for statistical test) between two individuals were enriched in allelic differences in the associated gene region as compared to the underlying allelic variation in those regions among all individuals (Figure 3A, left). For all genetic variant types (non-coding, synonymous exon, non-synonymous exon), we found statistical enrichment ($p < 0.05$, binomial distribution), suggesting the PERC, in part, depends on the genotype. Further, we filtered out reactions whose PERCs had high intra-individual differences longitudinally within the same individual (Supplemental Experimental Procedures). The filtering further increased enrichment levels. We next determined whether differences in alleles between individuals were enriched in ensembles of PERCs that were significantly different (Figure 3A, right). This test is more stringent as most genetic variants are not functional. Enrichment levels were much lower, while non-synonymous and non-coding genes were not significantly enriched with PERC differences. Filtering the same reactions with high intra-individual differences determined from the longitudinal controls from this analysis, all variant types were statistically enriched.

We then compared significantly different intracellular erythrocyte metabolite levels ($p < 0.05$, ANOVA, Bonferroni corrected) with genetic variants in gene regions associated with metabolic reactions producing or consuming the metabolite. Notably, significantly different metabolite levels were not enriched in differences of alleles ($p > 0.05$, binomial distribution) (Figure 3A, left), and differences in alleles also were not enriched in significantly different metabolite levels (Figure 3A, right). This is most likely because there is not a 1:1 relationship between metabolites and reactions. To account for this discrepancy, we separated the enrichment results into the single reactions associated with the significantly different metabolite. Next we chose the reaction that had the best enrichment with the associated metabolite to create a 1:1 relationship that represents the best-case scenario. Still, the metabolite differences were not as significantly enriched in differences of alleles as the ensemble of PERCs (Figure 3A, left), and allelic differences in selected genes were not significantly enriched in differences of metabolite concentration (Figure 3A, right). This result may be attributable to the following: (1) metabolite levels and enzyme activity are indirectly linked, as metabolite levels can vary in a system while the enzyme activity or concentration stays constant; and (2) cellular systems often modulate enzyme levels and activity to maintain metabolite levels.

To further characterize the relationship of PERCs and genetic variants, we found a significant correlation between the accuracy of predicting differences in ensembles of PERCs with the minor allele frequency (MAF, based on the 1000 Genomes Project) of the variant (Figure 3B). This result is consistent with the notion that functional variants are more likely to be rare (Cargill et al., 1999).

Temporal Decomposition

In addition to statistical analyses, personalized kinetic models can be analyzed for mechanistic-based physiological insights

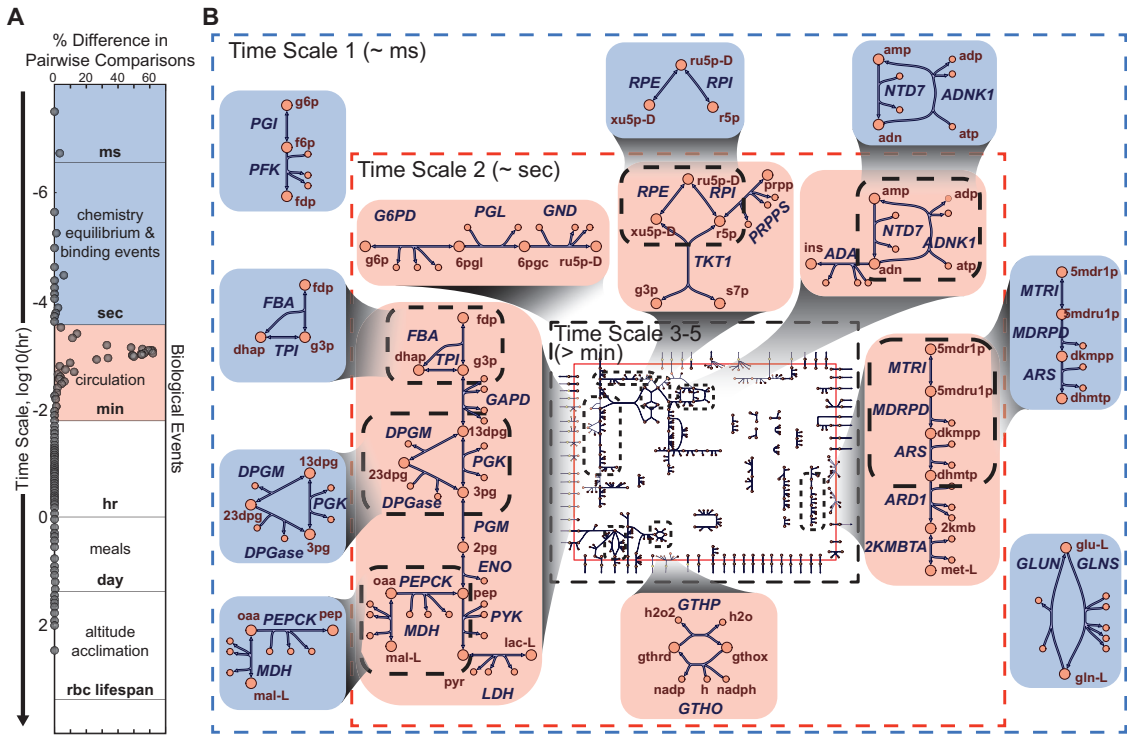


Figure 4. Temporal Decomposition of the Models Elucidated Personalized Systemic Variations at Specific Timescales

(A) The majority of differences in timescales between different individuals occurred on the order of seconds. This variability in dynamics coincides with the physiological timescales of blood circulation (~1 min).

(B) Metabolic reactions were aggregated based on the timescales of their coordinated activity. The first two timescales are shown. For example, aldolase (FBA) and triosephosphate isomerase (TPI) are fast reactions, operating near equilibrium, and coordinated within the first timescale. These two reactions then coordinate with the rest of lower glycolysis in the second timescale. Many of the important metabolic pathways for erythrocyte function aggregate on this time scale (~seconds).

that are not possible with genome-wide association studies. Specifically, we analyzed the individual variability in network dynamic responses by assessing the properties of the personalized Jacobian matrices (Palsson et al., 1987). The Jacobian matrix describes network dynamics around a particular homeostatic reference state (Heinrich et al., 1977; Jamshidi and Palsson, 2008b; Kauffman et al., 2002). Decomposing the Jacobian matrix yields the independent dynamic drivers, or modes, of the network, as well as the time constants under which such drivers operate (Jamshidi and Palsson, 2009; Experimental Procedures).

To assess variations in the dynamics across the entire network, we determined which time constants were significantly different between the ensemble models of individuals ($p < 0.05$, empirical test). Interestingly, almost all pair-wise differences in the time constants occurred between response times of 1 s and 1 min (Figure 4A), which coincide with the average capillary transit time (~1 s) and average erythrocyte circulation time (~1 min), suggesting that personalized metabolic functions may affect the oxygenation and deoxygenation processes. The median effect size of time constant differences was 14.5%, with effect sizes as large as 150%. Repeating the temporal decomposition analysis with the longitudinal controls, the results did not have the same specific difference as seen between individuals, suggesting the results are not an artifact of the approach (Supplemental Experimental Procedures; Figure S3).

Next we identified the dominant dynamic features of modes operating on different timescales (millisecond, second, minute, hour, and day) to ascertain which metabolic fluxes underlie the individual variation of the time constants. A previously published method (Jamshidi and Palsson, 2008b) was applied to determine when individual fluxes coordinate their activities on the separate timescales based on the modes. Reaction flux aggregation of the first two timescales are illustrated in Figure 4B. The first timescale (~millisecond) is the coordination of very fast reactions near equilibrium, such as isomerases (phosphoglucose isomerase, triosephosphate isomerase, and ribose-5-phosphate isomerase). The second timescale (~second), where most individual variation was detected, is characterized by the coordination of fluxes into parts of key erythrocyte pathways, including lower glycolysis, the oxidative pentose phosphate pathway, and glutathione recycling.

Dynamic Simulations for Understanding Ribavirin Susceptibility

Personalized kinetic models can be used to assess the systemic responses to perturbations through dynamic simulations (Figure 5A). We used simulations to study individual responses to ribavirin (RBV), an anti-viral drug used for the hepatitis C virus.

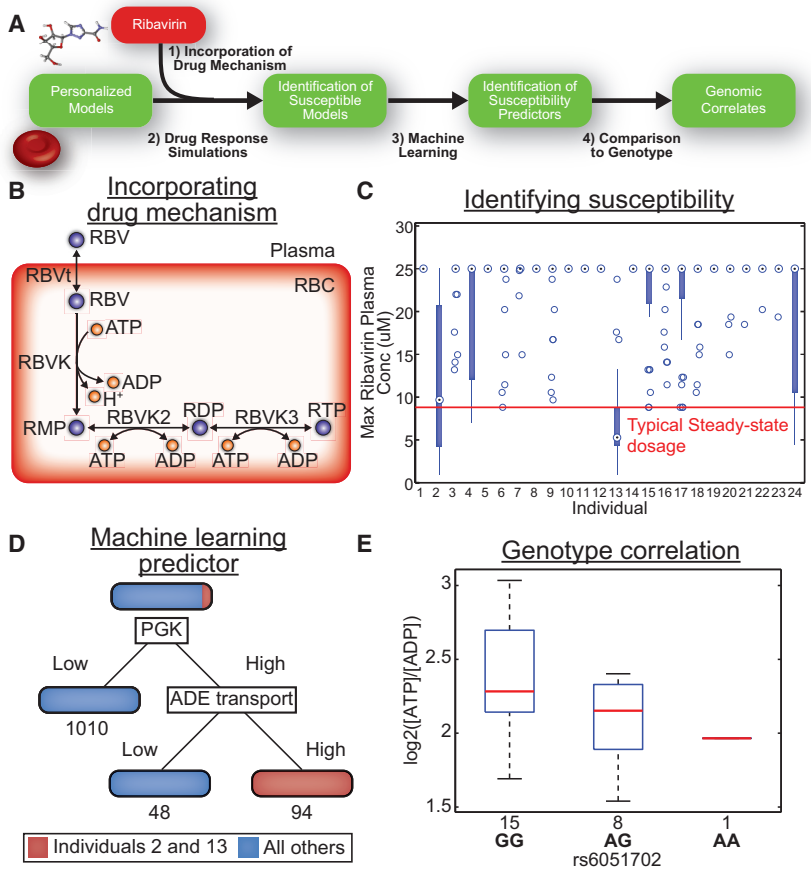


Figure 5. Dynamic Simulations of the Personalized Models Generated Novel Predictions about Susceptibility to Ribavirin-Induced Hemolytic Anemia

(A) Workflow for personalized drug exposure, identification of susceptible models, and predictors of susceptibility are shown.

(B) The ribavirin (RBV) metabolism pathway was added to each model in order to simulate the RBV side effect mechanism.

(C) The ensemble models were iteratively simulated at increasing doses of RBV to determine the plasma concentration at which the models of an individual could no longer return to a homeostatic state.

(D) A robust classification tree was constructed that predisposes whether or not models return to a steady state. High PERCs for phosphoglycerate kinase (PGK) and adenine transport were identified as factors of susceptibility.

(E) The ATP/ADP ratio in erythrocytes is correlated ($p = 0.0306$, adjusted slope of regression model $\beta = -0.28305$, log2 transformed) with the highest associated variant (rs6051702) of ITPA protection of RBV-induced hemolytic anemia.

A common side effect of RBV is dose-dependent non-immune-mediated hemolytic anemia (Russmann et al., 2006). The mechanism generally accepted for this side effect is that RBV is phosphorylated in the erythrocyte, but cannot be dephosphorylated (Figure 5B). Lack of dephosphorylation causes the accumulation of >1 mM RBV and its phosphorylated forms (RBV monophosphate, RBV diphosphate, and RBV triphosphate) in erythrocytes (Homma et al., 2004), sequestering >2 mM inorganic phosphate. Curiously, individuals with inosine triphosphatase (ITPA) deficiency, the inability to hydrolase inosine triphosphate to inosine monophosphate, are protected from this side effect (Fellay et al., 2010). ITPA deficiency is clinically benign but is marked with a buildup of inosine triphosphate in erythrocytes (van Waeg et al., 1988).

RBV phosphorylation kinetic expressions were added to the models (Supplemental Experimental Procedures). Dynamic simulations indicated significant decreases in ATP levels and flux through the sodium-potassium ATP pump, while there were no significant changes in glutathione levels. These simulated dynamic changes were qualitatively concordant with known changes found in RBV stimulation in vitro (De Franceschi et al., 2000). Using dynamic simulations, we identified the plasma RBV concentrations (up to three times the typical plasma RBV concentration; Brochot et al., 2010), of which each model was unable to return to a steady state due to drug exposure (Figure 5C). Most individuals (22 of 24) were non-responders, as the ensemble models for those individuals tolerated RBV. However, for two of the 24 individuals examined, a large proportion of their ensemble models were unable to return to a homeostatic state. The simulated side effect incidence rate (2/24, 8.3%) is concordant with the percentage of RBV patients (n = 1,184) during Phase III clinical trials (Devine et al., 2001) that had RBV dosage reduced (7.4%), discontinued (0.3%), or that received transfusions (0.1%). Further, eight individuals were heterozygous and one individual was homozygous for the variant (rs6051702; Fellay et al., 2010) highest associated with ITPA protection against RBV-induced hemolytic anemia. All nine of these individuals were in the non-responder group.

To determine if there was an underlying predisposition in models that were susceptible to RBV-induced hemolytic anemia, we used classification trees with the ensembles of PERCs as predictors to classify non-responders versus responders (Figure 5D). The classifier indicated that high PERCs for phosphoglycerate kinase (PGK, the first ATP-generating reaction in glycolysis) and adenine transport are a potential reason for RBV-induced hemolytic anemia. Previous clinical studies have identified individuals with hemolytic anemia that had high glycolytic rates and high adenine catabolism (Mills et al., 1968). A high glycolytic rate and elevated adenine catabolism enable quicker access of RBV to the ATP pool, thus increasing the RBV phosphorylation rate. Further, we modified the metabolite concentrations of the susceptible models of individuals 2 and 13 to lower the PERCs for PGK and adenine transport and were able to maintain model homeostasis with normal dosages of RBV (Supplemental Experimental Procedures; Figure S5). Thus, predisposition to RBV-induced hemolytic anemia may be due to an inability to adapt to high rates of RBV phosphorylation in certain individuals.

We next determined if protection by ITPA deficiency is related to the identified classifiers. One of the predisposition factors, the PERC of PGK (and its associated high glycolytic rate), is dependent on a high ATP/ADP ratio. ITPA deficiency, which leads to a buildup of ITP in erythrocytes, is in effect sequestering phosphate from the total phosphate pool. Changing the phosphate pool has been shown to affect the ATP/ADP ratio (Minakami and Yoshikawa, 1966). We found that the ATP/ADP ratio was significantly associated ($p = 0.0306$) with rs6051702 (Figure 5C). Thus, lowered ATP/ADP ratio due to a smaller phosphate pool may play a role in ITPA deficiency protection from RBV-induced hemolytic anemia.

DISCUSSION

We developed personalized whole-cell kinetic models of erythrocyte metabolism for 24 healthy individuals using a scalable, metabolomic-driven workflow. The metabolic network structure provides the non-linear relationships between metabolites and enzymes in the form of reaction rate laws. Thus, personal metabolite measurements can be interpreted mechanistically to deduce changes of functional importance, such as enzyme activities and cellular dynamics. The approach can be implemented for any cell type for which metabolomic data can be procured. To our knowledge, a unique feature of this work is that it represents a bottom-up mechanistic approach to bridge the variation in dynamic responses of human cellular metabolism with the underlying genotype, utilizing high-throughput data. Implementation of the MASS approach for building kinetic models results in predictive models that have implications for accelerating discoveries in understanding human metabolic variation.

Our analysis shows that PERCs, an approximation for enzyme activity and abundance, are more associated with allelic variants than metabolite concentrations. PERCs also have larger effect sizes, making them more sensitive for detecting subtleties in human metabolic variation than metabolite quantitative trait loci (Suhre and Gieger, 2012; Suhre et al., 2011). Owing to our sample size, we only globally studied whether changes in variants within a gene region of an enzyme were enriched with changes in the value of the enzyme's PERC. We anticipate that scale-up to sample sizes more typical of genome-wide association studies will provide sufficient statistical power to allow PERCs to be associated with genetic loci, leading to discoveries in both how coding variants affect enzyme kinetic properties and how non-coding variants across the genome affect enzyme abundance.

Model-driven analysis led us to propose a mechanism for predisposition to RBV-induced hemolytic anemia. The mechanism is consistent with the well-known protective properties of ITPA deficiency. An alternative mechanism has been proposed (Hitomi et al., 2011) but with notable shortcomings (Supplemental Experimental Procedures). Follow-up of our mechanism with larger populations is warranted to determine if the ATP/ADP ratio is in fact a major determinant of susceptibility. Interestingly, a recent study (Zhao et al., 2013) analyzed the FDA Adverse Event Reporting System and identified drug combinations (Drug A + Drug B) that had lower side effect rates than Drug A alone. Of the detected 19,000 combinations, two were related to RBV + Drug B (either ketoprofen or prednisolone), which lowered the rate of occurrence of hemolytic anemia as compared to RBV alone. Ketoprofen is a known weak inhibitor of bacterial glycolytic metabolism (Belli et al., 1995), concordant with our proposed mechanism. Our results highlight the utility of kinetic models for pharmacodynamics. Although the hypothesis can be tested with metabolomic data, the insight and generation of the hypothesis required the integration of metabolomic and genotyping data with kinetic modeling.

In contrast to parameter-fitting-based approaches, we apply conservation principles and thermodynamic constraints to the biochemical network and calculate parameters (PERCs) based on the resulting feasible points within the flux solution space. Our findings demonstrate the feasibility and utility of these models for studying inter-individual variations in physiological dynamics and drug responses. We anticipate that mechanistically driven studies that incorporate multiple types of omic data will become effective means to understanding human variation and its implications.

EXPERIMENTAL PROCEDURES

Study Sample

University of California, San Diego (UCSD) institutional review board approval was obtained (Project 111100X) permitting the acquisition and subsequent analysis of blood from healthy volunteers via peripheral venipuncture. The 24 individuals (16 male, 8 female) who identified themselves as healthy, without any known chronic medical conditions, provided written informed consent and provided venous blood samples for plasma and erythrocyte metabolite profiling as well as SNP genotyping in a morning fasted state (a minimum of 8 hr). Four of the individuals (individuals 1–4) were sampled for metabolite profiling four times across 6 weeks as longitudinal controls (weeks 0, 1, 4, and 6). Their final time point was used in the comparisons against the other 20 individuals.

Sample Preparation

Two vials of blood (~8 ml) were drawn from each individual. The first vial (included in PAXgene blood DNA kit, QIAGEN) subsequently was used for DNA extraction and SNP genotyping. The Illumina Infinium HumanOmniExpressExome BeadChip was used for SNP genotyping. The UCSD IGM Genomics Center (http://igm.ucsd.edu/core-services/) completed analyses and made SNP calls. The second vial, containing heparin anticoagulant, was immediately placed on ice. The vial was split for plasma and intracellular erythrocyte metabolite extraction. Erythrocyte samples were centrifuged (1,000 × *g*, 2 min, 4°C). Then, 25 μl packed red cells were aliquoted with 25 μl internal standards (next section) and 450 μl −80°C methanol (80%). Tubes were immersed in liquid nitrogen until frozen to lyse cells. After thawing in a water bath at room temperature, tubes were centrifuged (15,800 × *g*, 15 min, 4°C). Supernatant was retained for metabolite profiling and stored at −80°C. Concurrently, plasma samples were centrifuged (3,000 × *g*, 20 min, 4°C). Then, 100 μl plasma was aliquoted with 35 μl internal standards and 365 μl −80°C methanol (80%). Tubes were centrifuged (15,800 × *g*, 15 min, 4°C). The supernatant was retained for metabolite profiling and stored at −80°C.

All samples were lyophilized and reconstituted in water and ran at the nominal concentration and at a 10-fold dilution to account for the large concentration range of intracellular metabolites. Metabolite samples were prepared in triplicate. Across all samples, 2.8% and 4.0% of the erythrocyte and plasma metabolomes, respectively, were not detected. These values were imputed using the weighted average of the five nearest neighbors (Tong, 2008). The metabolomic data are presented in Data S1. Some metabolites cannot be differentiated experimentally by global mass spectrometry (e.g., 2pg and 3pg) and are measured in pools. To not introduce additional variation in the personalized models, the ratios of the pooled metabolites were kept constant in the personalized models as the baseline model, just the total size of the pool was varied between individuals.

Metabolite Profiling: Standards, Instrumentation, Acquisition, and Quantification

Standards were purchased from Sigma-Aldrich or Santa Cruz Biotechnology. Liquid chromatography-mass spectrometry (LC-MS) reagents were purchased from Honeywell Burdick & Jackson. Metabolically labeled internal standards were synthesized as described previously (McCloskey et al., 2014). Calibrator standards were generated from stock solutions freshly prepared or kept in −80°C for no longer than 3 days. Calibrator standards were combined into mixtures and aliquoted, lyophilized, and stored at −80°C. Aliquots were reconstituted in water and serially diluted to generate a calibration curve that spanned the lower and upper limits of detection for each compound.

An XSELECT HSS XP 150 mm × 2.1 mm × 2.5 μm (Waters Corporation) with a UFLC XR HPLC (Shimadzu Scientific Instruments) was used for chromatographic separation. Mobile phase A was composed of 10 mM tributylamine (TBA), 10 mM acetic acid (pH 6.86), 5% methanol, and 2% 2-propanol; mobile phase B was 2-propanol. Oven temperature was 40°C. The chromatographic conditions were as follows: 0, 0, 0.4; 5, 0, 0.4; 9, 2, 0.4; 9.5, 6, 0.4; 11.5, 6, 0.4; 12, 11, 0.4; 13.5, 11, 0.4; 15.5, 28, 0.4; 16.5, 53, 0.15; 22.5, 53, 0.15; 23, 0, 0.15; 27, 0, 0.4; and 33, 0, 0.4; (total time [min], eluent B [vol %], flow rate [ml/min]). The autosampler temperature was 10°C and the injection volume was 10 μl with full loop injection. An AB SCIEX Qtrap 5500 mass spectrometer (AB SCIEX) was operated in negative mode. Electrospray ionization parameters were optimized for 0.4 ml/min flow rate and were as follows: electrospray voltage of −4,500 V, temperature of 500°C, curtain gas of 40, CAD gas of 12, and gases 1 and 2 of 50 and 50 psi, respectively. Analyzer parameters were optimized for each compound using manual tuning. The instrument was mass calibrated with a mixture of polypropylene glycol (PPG) standards.

Samples were acquired using the scheduled multiple reaction monitoring (MRM) pro algorithm in Analyst 1.6.2. The acquisition method consisted of an MRM survey scan coupled to an information-dependent acquisition (IDA) consisting of an enhanced product ion (EPI) scan for compound identity confirmation. Samples were quantified using IDMS39, 40 with metabolically labeled internal standards and processed using Multiquant 2.1.1. Linear regressions for compound quantitation were based on peak height ratios and the logarithm of the concentration of calibrator concentrations from a minimum of four consecutive concentration ranges that showed minimal bias. A peak height greater than 1e3 ion counts and a signal-to-noise ratio greater than 20 were used to define the lower limit of quantitation (LLOQ). Quality controls and carryover checks were included with each batch. Due to the number of biological isomers, the integration of each compound was checked manually.

Constraint-Based Model Construction and Flux Calculations

The human erythrocyte metabolic model was updated using COBRA Toolbox v2 (Schellenberger et al., 2011) for MATLAB (MathWorks). Model changes included removing pathways whose majority of metabolites were not measured and adding reactions based on the additional scope from metabolomic measurements and a recently published proteomic dataset (D'Alessandro et al., 2010). Data S1 details specific changes to the model. Literature sources (Data S1) were used to set transport flux rates into and out of the model as well as flux rates of NADH and ATP usage, catalase activity, and flux ratio splits. Lower and upper bounds for these fluxes were set as the mean ± SD, where available. Otherwise, flux bounds were set as 75%–125% of the mean. MCMC sampling (Bordbar et al., 2010; Schellenberger et al., 2011) was used to determine candidate flux states or sample points; 5,000 sample points were computed and the mean was used as the initial flux state used to build the baseline model. Of the 5,000, 300 sample points were randomly chosen to build the ensemble models (Data S1).

Kinetic Model Construction

Kinetic model construction, temporal decomposition, and dynamic simulations were completed with the MASS Toolbox (http://opencobra.github.io/MASS-Toolbox/) for Mathematica (Wolfram). The MASS modeling approach and the workflow for building the models have been previously documented (Jamshidi and Palsson, 2008a; Jamshidi and Palsson, 2010). Initial metabolite concentrations were set based on literature sources. Plasma concentrations were assumed to be constant. Concentrations for some metabolites were not available and were set based on MCMC concentration sampling of thermodynamically feasible ranges (Equation 1), where S is the stoichiometric matrix. The known concentrations were fixed and the remaining concentrations were calculated. The highest probability concentrations were chosen.

$$\begin{aligned} &S^T \cdot \log(x) \le \log(K_{eq}) \text{if } v>0 \\ &S^T \cdot \log(x) \ge \log(K_{eq}) \text{if } v<0 \\ &lb = ub = x_{known} \end{aligned} \qquad \text{(Equation 1)}$$

Equilibrium constants were set based on experimental measurements, where available, or computational methods (Noor et al., 2012; http://equilibrator.weizmann.ac.il). Equilibrium constants for transport reactions were set based on a previous method (Jol et al., 2010). Literature sources for metabolite concentrations and equilibrium constants are provided in Data S1. Model stability was assessed by calculating the Jacobian matrix and checking that eigenvalues were negative or zero within a tolerance of 10^{-10}. To achieve linear stability of the baseline model, the computationally determined equilibrium constants were varied up to three orders of magnitude and the eigenvalues of the Jacobian matrix were re-calculated to determine the minimal adjustment needed for a stable kinetic model.

The internal standards were used to relatively quantify the plasma and erythrocyte concentrations, and the mean concentration of a metabolite across all samples was assumed to be the baseline concentration. Measured plasma and erythrocyte concentrations were minimally adjusted using a quadratic programming (QP) problem to ensure thermodynamic consistency (Equation 2). Γ is the mass action ratio; x_0 and Γ_0 are the baseline model concentrations and mass action ratio; ε was set to 10% of the difference of Γ and K_{eq}. This ensures that the calculated PERC in later steps is not infinity or negative infinity.

$$\begin{aligned} &\min(\log(x/x_0))^2 + (\log(\Gamma/\Gamma_0))^2 \\ &S^T \cdot \log(x) \le \log(K_{eq}) - \varepsilon \text{ if } v>0 \\ &S^T \cdot \log(x) \ge \log(K_{eq}) + \varepsilon \text{ if } v<0 \\ &\log(0.5) \le \log(x/x_0) \le \log(1.5) \end{aligned} \qquad \text{(Equation 2)}$$

After integration of personal metabolite data, the measured metabolite concentrations were varied up to an order of magnitude and the eigenvalues of the Jacobian were re-calculatd to determine the minimal adjustment needed for a stable kinetic model. Once stabilized, 300 ensemble models were built for each individual model and the baseline model using MCMC sample points. The models were filtered based on their ability to simulate to a steady state due to a perturbation by decreasing the ATP concentration by 5% and increasing the ADP concentration by 5%. The number of stable models varied for each individual from 62 to 287 models, with 48 flux states having a stable kinetic model for all individuals and the baseline. The 48 models for each individual were utilized in the study.

The baseline kinetic models built by the MASS and enzymatic rate law approaches are available in Data S2, Data S3, and BioModels (http://www.ebi.ac.uk/biomodels-main/). The construction of the enzymatic rate law model is presented in the Supplemental Experimental Procedures and Data S1.

Statistical Analyses for PERC Comparison with Allele Differences

When comparing differences of PERCs and metabolites to differences in alleles, transporter reactions were excluded, as stoichiometries of particular transporters may not be known. PERCs were considered to be significantly different between individuals if the distribution of PERCs from the 48 ensemble models had no overlap and had a median effect size of at least a factor of 2. Instead of using the overlap of distributions, had we used the non-parametric Wilcoxon rank-sum test, the two-sided p value would be p = 3.23e−17, which is well below the Bonferroni adjusted p value for the number of tests done (p = 2.16e−7).

To determine the association of the log2-transformed ATP/ADP ratio with rs6051702, a linear regression model was used, accounting for sex of the volunteer as a potential confounder. Sex was included as a potential confounder as the erythrocyte Na-K ATPase has different kinetic properties in males and females (Lasker et al., 1985).

Temporal Decomposition

There is a dual set of Jacobian matrices (a flux component, J_v, and a metabolite component, J_x) and they describe the network dynamics near the homeostatic reference state (Jamshidi and Palsson, 2009). Using a similarity transformation ($J = M\Lambda M^{-1}$), the Jacobian matrix can be decomposed into its

independent dynamic components (modal matrix, M^{-1}) and their corresponding eigenvalues (Λ), which are unique for each dual Jacobian pair (Jamshidi and Palsson, 2008a). The modes are expressed either as fluxes or metabolites, depending on which Jacobian matrix is decomposed. The eigenvalues derived from the similarity analysis of the dual Jacobian matrices are the same. The time constants for each mode are the negative reciprocal of the corresponding eigenvalue. Decomposing the ensemble metabolic models, there were 105 distinct modes, or dynamically independent groups of metabolites (or fluxes), in all of the personalized models. To determine the aggregation of reaction fluxes across different timescales, a previously developed method was used (Jamshidi and Palsson, 2008b) with a correlation coefficient threshold of 0.85.

Classification Tree

To determine potential disposition factors to RBV susceptibility, we classified the two individuals that were responders to the drug versus the non-responders, using PERCs as the predictive variables. One-third of the ensemble models (split evenly between responders and non-responders) were left out of the training to be our test set. Training and validation were completed on the remaining two-thirds of models using a 10-fold cross-validation. After building the classifier, the predictors were used for the test set. We repeated this process 1,000 times, removing different thirds of the models. The classification tree was very robust, with the average misclassification percentage being 0.49% and the maximum misclassification percentage in any iteration being 1.67%. For all 1,000 iterations, the two main predictors were PGK and adenine transfer. The final classification tree was derived from all samples.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, five figures, and three data files and can be found with this article online at http://dx.doi.org/10.1016/j.cels.2015.10.003.

AUTHOR CONTRIBUTIONS

A.B. and N.J. prepared metabolite and DNA extracts. D.M. completed mass spectrometry and related analyses. A.B. built the models, completed all analyses, and wrote the manuscript. D.C.Z. and N.S. aided in kinetic modeling. A.B., N.J., and B.O.P. conceived and designed the study. A.B., N.J., and B.O.P. wrote the paper. All authors edited and approved the final manuscript.

ACKNOWLEDGMENTS

We would like to thank the participants of this project. We also thank N.E. Lewis, J.M. Monk, A. Ebrahim, E. Ruppin, and N.D. Price for valuable discussions, as well as A. Draeger for standardizing models for the Systems Biology Markup Language (SBML). Whole-genome genotyping was conducted at the IGM Genomics Center, UCSD, La Jolla, CA. This work was supported by NIH grant GM068837 and the Novo Nordisk Foundation. N.J. and B.O.P. have an issued patent for large-scale omic data-driven kinetic modeling, as applied to studying individuals.

Received: May 27, 2015
Revised: August 13, 2015
Accepted: October 7, 2015
Published: October 28, 2015

REFERENCES

Agren, R., Mardinoglu, A., Asplund, A., Kampf, C., Uhlen, M., and Nielsen, J. (2014). Identification of anticancer drugs for hepatocellular carcinoma through personalized genome-scale metabolic modeling. Mol. Syst. Biol. *10*, 721.

Beckonert, O., Keun, H.C., Ebbels, T.M., Bundy, J., Holmes, E., Lindon, J.C., and Nicholson, J.K. (2007). Metabolic profiling, metabolomic and metabonomic procedures for NMR spectroscopy of urine, plasma, serum and tissue extracts. Nat. Protoc. *2*, 2692–2703.

Belli, W.A., Buckley, D.H., and Marquis, R.E. (1995). Weak acid effects and fluoride inhibition of glycolysis by Streptococcus mutans GS-5. Can. J. Microbiol. *41*, 785–791.

Bordbar, A., Lewis, N.E., Schellenberger, J., Palsson, B.O., and Jamshidi, N. (2010). Insight into human alveolar macrophage and M. tuberculosis interactions via metabolic reconstructions. Mol. Syst. Biol. *6*, 422.

Bordbar, A., Jamshidi, N., and Palsson, B.O. (2011). iAB-RBC-283: A proteomically derived knowledge-base of erythrocyte metabolism that can be used to simulate its physiological and patho-physiological states. BMC Syst. Biol. *5*, 110.

Bordbar, A., Monk, J.M., King, Z.A., and Palsson, B.O. (2014). Constraint-based models predict metabolic and associated cellular functions. Nat. Rev. Genet. *15*, 107–120.

Brochot, E., Castelain, S., Duverlie, G., Capron, D., Nguyen-Khac, E., and François, C. (2010). Ribavirin monitoring in chronic hepatitis C therapy: anaemia versus efficacy. Antivir. Ther. (Lond.) *15*, 687–695.

Cargill, M., Altshuler, D., Ireland, J., Sklar, P., Ardlie, K., Patil, N., Shaw, N., Lane, C.R., Lim, E.P., Kalyanaraman, N., et al. (1999). Characterization of single-nucleotide polymorphisms in coding regions of human genes. Nat. Genet. *22*, 231–238.

Chakrabarti, A., Miskovic, L., Soh, K.C., and Hatzimanikatis, V. (2013). Towards kinetic modeling of genome-scale metabolic networks without sacrificing stoichiometric, thermodynamic and physiological constraints. Biotechnol. J. *8*, 1043–1057.

Chen, R., Mias, G.I., Li-Pook-Than, J., Jiang, L., Lam, H.Y., Chen, R., Miriami, E., Karczewski, K.J., Hariharan, M., Dewey, F.E., et al. (2012). Personal omics profiling reveals dynamic molecular and medical phenotypes. Cell *148*, 1293–1307.

D'Alessandro, A., Righetti, P.G., and Zolla, L. (2010). The red blood cell proteome and interactome: an update. J. Proteome Res. *9*, 144–163.

Danino, T., Mondragón-Palomino, O., Tsimring, L., and Hasty, J. (2010). A synchronized quorum of genetic clocks. Nature *463*, 326–330.

De Franceschi, L., Fattovich, G., Turrini, F., Ayi, K., Brugnara, C., Manzato, F., Noventa, F., Stanzial, A.M., Solero, P., and Corrocher, R. (2000). Hemolytic anemia induced by ribavirin therapy in patients with chronic hepatitis C virus infection: role of membrane oxidative damage. Hepatology *31*, 997–1004.

Devine, E.B., Kowdley, K.V., Veenstra, D.L., and Sullivan, S.D. (2001). Management strategies for ribavirin-induced hemolytic anemia in the treatment of hepatitis C: clinical and economic implications. Value Health *4*, 376–384.

Duarte, N.C., Becker, S.A., Jamshidi, N., Thiele, I., Mo, M.L., Vo, T.D., Srivas, R., and Palsson, B.O. (2007). Global reconstruction of the human metabolic network based on genomic and bibliomic data. Proc. Natl. Acad. Sci. USA *104*, 1777–1782.

Feist, A.M., Herrgard, M.J., Thiele, I., Reed, J.L., and Palsson, B.O. (2009). Reconstruction of biochemical networks in microorganisms. Nat. Rev. Microbiol. *7*, 129–143.

Fellay, J., Thompson, A.J., Ge, D., Gumbs, C.E., Urban, T.J., Shianna, K.V., Little, L.D., Qiu, P., Bertelsen, A.H., Watson, M., et al. (2010). ITPA gene variants protect against anaemia in patients treated for chronic hepatitis C. Nature *464*, 405–408.

Heinrich, R., Rapoport, S.M., and Rapoport, T.A. (1977). Metabolic regulation and mathematical models. Prog. Biophys. Mol. Biol. *32*, 1–82.

Hitomi, Y., Cirulli, E.T., Fellay, J., McHutchison, J.G., Thompson, A.J., Gumbs, C.E., Shianna, K.V., Urban, T.J., and Goldstein, D.B. (2011). Inosine triphosphate protects against ribavirin-induced adenosine triphosphate loss by adenylosuccinate synthase function. Gastroenterology *140*, 1314–1321.

Holmes, E., Wilson, I.D., and Nicholson, J.K. (2008). Metabolic phenotyping in health and disease. Cell *134*, 714–717.

Homma, M., Matsuzaki, Y., Inoue, Y., Shibata, M., Mitamura, K., Tanaka, N., and Kohda, Y. (2004). Marked elevation of erythrocyte ribavirin levels in interferon and ribavirin-induced anemia. Clin. Gastroenterol. Hepatol. *2*, 337–339.

Jamshidi, N., and Palsson, B.O. (2006). Systems biology of SNPs. Mol. Syst. Biol. *2*, 38.

Jamshidi, N., and Palsson, B.O. (2008a). Formulating genome-scale kinetic models in the post-genome era. Mol. Syst. Biol. *4*, 171.

Jamshidi, N., and Palsson, B.O. (2008b). Top-down analysis of temporal hierarchy in biochemical reaction networks. PLoS Comput. Biol. *4*, e1000177.

Jamshidi, N., and Palsson, B.O. (2009). Flux-concentration duality in dynamic nonequilibrium biological networks. Biophys. J. *97*, L11–L13.

Jamshidi, N., and Palsson, B.O. (2010). Mass action stoichiometric simulation models: incorporating kinetics and regulation into stoichiometric models. Biophys. J. *98*, 175–185.

Jamshidi, N., Wiback, S.J., and Palsson B, B.Ø. (2002). In silico model-driven assessment of the effects of single nucleotide polymorphisms (SNPs) on human red blood cell metabolism. Genome Res. *12*, 1687–1692.

Jamshidi, N., Miller, F.J., Mandel, J., Evans, T., and Kuo, M.D. (2011). Individualized therapy of HHT driven by network analysis of metabolomic profiles. BMC Syst. Biol. *5*, 200.

Jol, S.J., Kümmel, A., Hatzimanikatis, V., Beard, D.A., and Heinemann, M. (2010). Thermodynamic calculations for biochemical transport and reaction processes in metabolic networks. Biophys. J. *99*, 3139–3144.

Kauffman, K.J., Pajerowski, J.D., Jamshidi, N., Palsson, B.O., and Edwards, J.S. (2002). Description and analysis of metabolic connectivity and dynamics in the human red blood cell. Biophys. J. *83*, 646–662.

Khodayari, A., Zomorrodi, A.R., Liao, J.C., and Maranas, C.D. (2014). A kinetic model of Escherichia coli core metabolism satisfying multiple sets of mutant flux data. Metab. Eng. *25*, 50–62.

Lanpher, B., Brunetti-Pierri, N., and Lee, B. (2006). Inborn errors of metabolism: the flux from Mendelian to complex diseases. Nat. Rev. Genet. *7*, 449–460.

Lasker, N., Hopp, L., Grossman, S., Bamforth, R., and Aviv, A. (1985). Race and sex differences in erythrocyte Na+, K+, and Na+-K+-adenosine triphosphatase. J. Clin. Invest. *75*, 1813–1820.

McCloskey, D., Gangoiti, J.A., King, Z.A., Naviaux, R.K., Barshop, B.A., Palsson, B.O., and Feist, A.M. (2014). A model-driven quantitative metabolomics analysis of aerobic and anaerobic metabolism in E. coli K-12 MG1655 that is biochemically and thermodynamically consistent. Biotechnol. Bioeng. *111*, 803–815.

Mills, G.C., Levin, W.C., and Alperin, J.B. (1968). Hemolytic anemia associated with low erythrocyte ATP. Blood *32*, 15–32.

Minakami, S., and Yoshikawa, H. (1966). Studies on erythrocyte glycolysis. 3. The effects of active cation transport, pH and inorganic phosphate concentration on erythrocyte glycolysis. J. Biochem. *59*, 145–150.

Noor, E., Bar-Even, A., Flamholz, A., Lubling, Y., Davidi, D., and Milo, R. (2012). An integrated open framework for thermodynamics of reactions that combines accuracy and coverage. Bioinformatics *28*, 2037–2044.

O'Dea, E.L., Barken, D., Peralta, R.Q., Tran, K.T., Werner, S.L., Kearns, J.D., Levchenko, A., and Hoffmann, A. (2007). A homeostatic model of IkappaB metabolism to control constitutive NF-kappaB activity. Mol. Syst. Biol. *3*, 111.

Oberhardt, M.A., Palsson, B.O., and Papin, J.A. (2009). Applications of genome-scale metabolic reconstructions. Mol. Syst. Biol. *5*, 320.

Orth, J.D., Thiele, I., and Palsson, B.O. (2010). What is flux balance analysis? Nat. Biotechnol. *28*, 245–248.

Palsson, B.O. (2011). Systems Biology: Simulation of Dynamic Network States (Cambridge University Press).

Palsson, B.O., Joshi, A., and Ozturk, S.S. (1987). Reducing complexity in metabolic networks: making metabolic meshes manageable. Fed. Proc. *46*, 2485–2489.

Pollitt, R.J., Green, A., McCabe, C.J., Booth, A., Cooper, N.J., Leonard, J.V., Nicholl, J., Nicholson, P., Tunaley, J.R., and Virdi, N.K. (1997). Neonatal screening for inborn errors of metabolism: cost, yield and outcome. Health Technol. Assess. *1*, i–iv, 1–202.

Psychogios, N., Hau, D.D., Peng, J., Guo, A.C., Mandal, R., Bouatra, S., Sinelnikov, I., Krishnamurthy, R., Eisner, R., Gautam, B., et al. (2011). The human serum metabolome. PLoS ONE *6*, e16957.

Russmann, S., Grattagliano, I., Portincasa, P., Palmieri, V.O., and Palasciano, G. (2006). Ribavirin-induced anemia: mechanisms, risk factors and related targets for future research. Curr. Med. Chem. *13*, 3351–3357.

Saucerman, J.J., Healy, S.N., Belik, M.E., Puglisi, J.L., and McCulloch, A.D. (2004). Proarrhythmic consequences of a KCNQ1 AKAP-binding domain mutation: computational models of whole cells and heterogeneous tissue. Circ. Res. *95*, 1216–1224.

Schellenberger, J., Que, R., Fleming, R.M., Thiele, I., Orth, J.D., Feist, A.M., Zielinski, D.C., Bordbar, A., Lewis, N.E., Rahmanian, S., et al. (2011). Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox v2.0. Nat. Protoc. *6*, 1290–1307.

Shlomi, T., Cabili, M.N., and Ruppin, E. (2009). Predicting metabolic biomarkers of human inborn errors of metabolism. Mol. Syst. Biol. *5*, 263.

Stanford, N.J., Lubitz, T., Smallbone, K., Klipp, E., Mendes, P., and Liebermeister, W. (2013). Systematic construction of kinetic models from genome-scale metabolic networks. PLoS ONE *8*, e79195.

Suhre, K., and Gieger, C. (2012). Genetic variation in metabolic phenotypes: study designs and applications. Nat. Rev. Genet. *13*, 759–769.

Suhre, K., Shin, S.Y., Petersen, A.K., Mohney, R.P., Meredith, D., Wägele, B., Altmaier, E., Deloukas, P., Erdmann, J., Grundberg, E., et al.; CARDIoGRAM (2011). Human metabolic individuality in biomedical and pharmaceutical research. Nature *477*, 54–60.

Tan, Y., and Liao, J.C. (2012). Metabolic ensemble modeling for strain engineers. Biotechnol. J. *7*, 343–353.

Tautenhahn, R., Cho, K., Uritboonthai, W., Zhu, Z., Patti, G.J., and Siuzdak, G. (2012). An accelerated workflow for untargeted metabolomics using the METLIN database. Nat. Biotechnol. *30*, 826–828.

Teusink, B., Passarge, J., Reijenga, C.A., Esgalhado, E., van der Weijden, C.C., Schepper, M., Walsh, M.C., Bakker, B.M., van Dam, K., Westerhoff, H.V., and Snoep, J.L. (2000). Can yeast glycolysis be understood in terms of in vitro kinetics of the constituent enzymes? Testing biochemistry. Eur. J. Biochem. *267*, 5313–5329.

Tong, L.V. (2008). Development and application of mass spectrometry-based metabolomics methods for disease biomarker identification. PhD thesis (Cambridge: Massachusetts Institute of Technology).

van Waeg, G., Niklasson, F., Ericson, A., and de Verdier, C.H. (1988). Purine metabolism in normal and ITP-pyrophosphohydrolase-deficient human erythrocytes. Clin. Chim. Acta *171*, 279–292.

Wishart, D.S., Jewison, T., Guo, A.C., Wilson, M., Knox, C., Liu, Y., Djoumbou, Y., Mandal, R., Aziat, F., Dong, E., et al. (2013). HMDB 3.0–The Human Metabolome Database in 2013. Nucleic Acids Res. *41*, D801–D807.

Yizhak, K., Gaude, E., Le Dévédec, S., Waldman, Y.Y., Stein, G.Y., van de Water, B., Frezza, C., and Ruppin, E. (2014). Phenotype-based cell-specific metabolic modeling reveals metabolic liabilities of cancer. eLife *3*, e03641.

Zhao, S., Nishimura, T., Chen, Y., Azeloglu, E.U., Gottesman, O., Giannarelli, C., Zafar, M.U., Benard, L., Badimon, J.J., Hajjar, R.J., et al. (2013). Systems pharmacology of adverse event mitigation by drug combinations. Sci. Transl. Med. *5*, 206ra140.