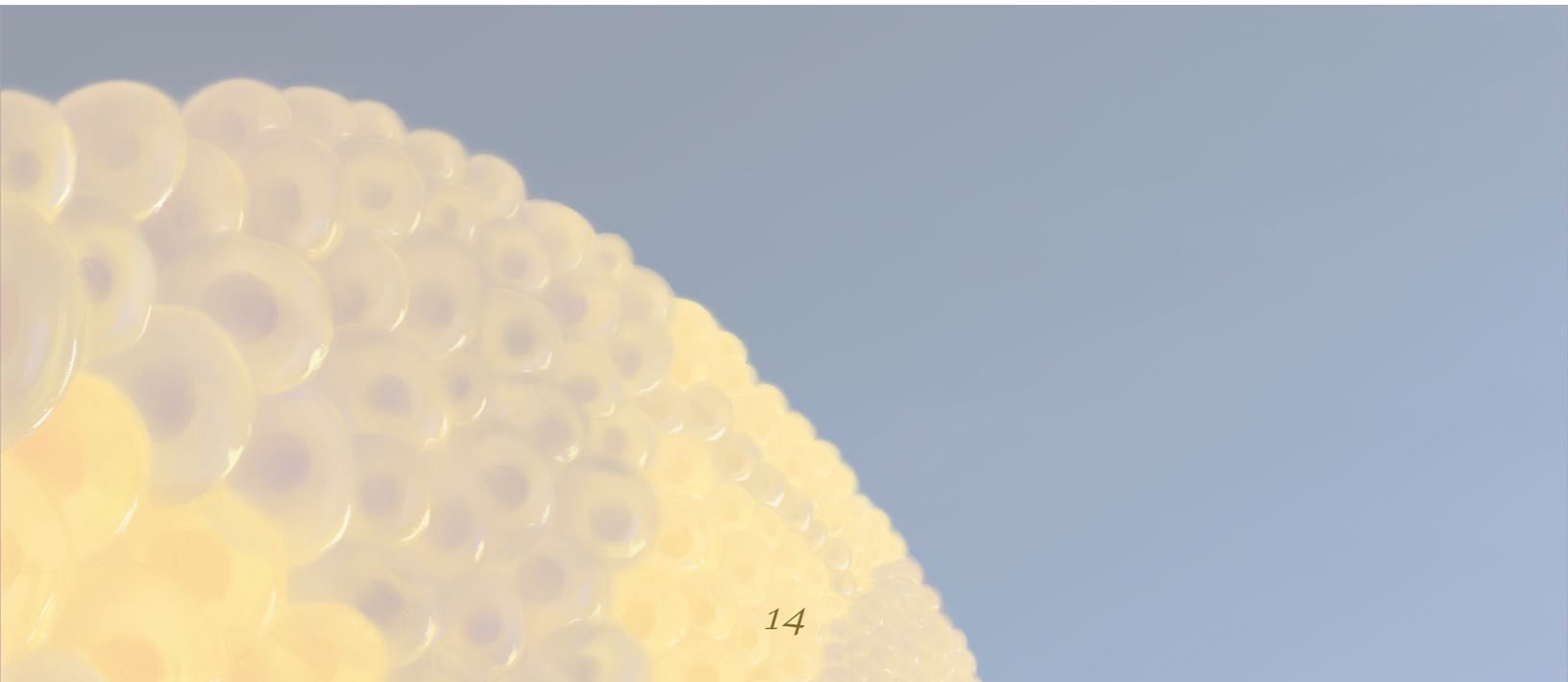
Using Precision Medicine to Predict Drug Efficacy and Toxicity

Paul W. Burridge, PhD

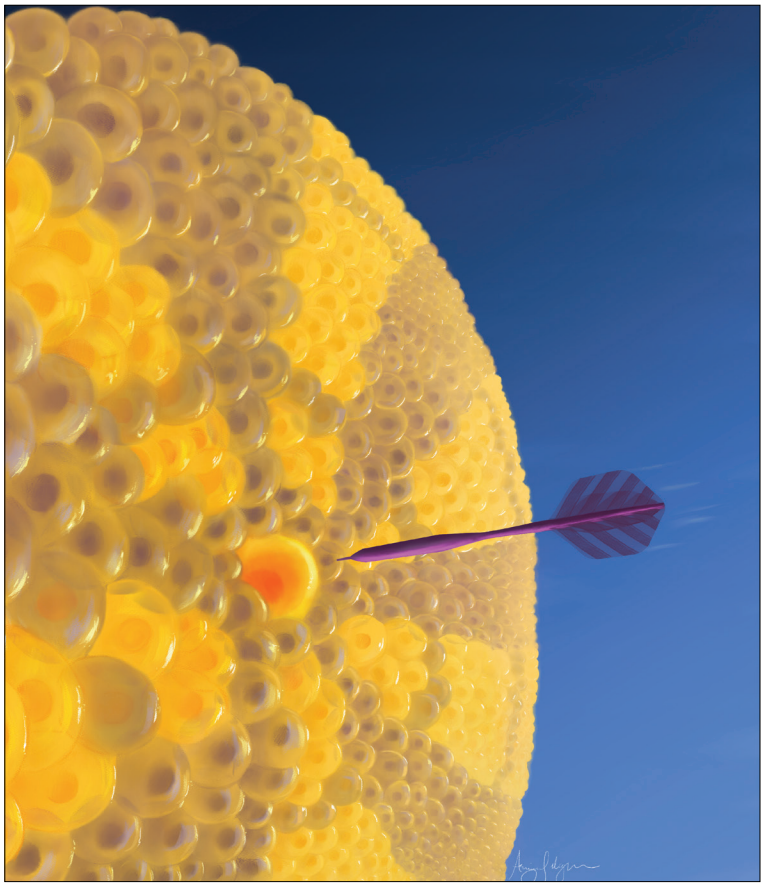
Department of Pharmacology, Center for Pharmacogenomics, Northwestern University Feinberg School of Medicine.

The precision medicine revolution is upon us. The ability to collect a variety of patient data, computationally combine and process it, and algorithmically pinpoint a disease management or preventative strategy will profoundly change patient treatment. In the future your doctors will not only know which diseases you are most likely to experience, they will also know how an illness will progress, which drugs should be used for treatment, when and in which combinations, and whether you are likely to suffer side-effects from these interventions.

The wealth of information furnished by a patient's genome makes genomic data a more valuable clinical tool than all other forms of patient information. With advances in DNA sequencing technology, and the cost of research-grade whole-genome sequencing (WGS) now only $600, there are no other precision medicine tools with such an extreme data-power-to-cost ratio. Genome sequencing is also a one-time cost, as your genome remains essentially the same throughout your life. Although the technology required to read a patient's three billion base-pairs of genomic information is now cheaper and more readily accessible than ever before, our knowledge of how to interpret and act upon this wellspring of data is still profoundly lacking.



The central aim of the field of pharmacogenomics is to investigate the relationship between patient drug response and the genome. The primary pharmacogenomics research method so far has been the genome-wide association study (GWAS), in which DNA is compared from patients who either do or do not have a specific positive or negative drug response. We can then locate single base-pair changes in DNA called single nucleotide polymorphisms (SNPs) that statistically correlate with this difference in drug response. This methodology has been used on a large scale by companies such as 23andMe, where they have collected DNA from millions of subjects and asked them to complete questionnaires to self-identify phenotypes. Conveniently, these participants can be contacted repeatedly to ask questions further clarifying their phenotypes or to collect responses to additional phenotype questions that arise once genotype data have been analyzed. Of course, phenotype-genotype data collected by private companies is proprietary and is therefore not a resource available to the global research community.



2018 ©Arianna Selagea
(UIC Biomedical Visualization graduate student)

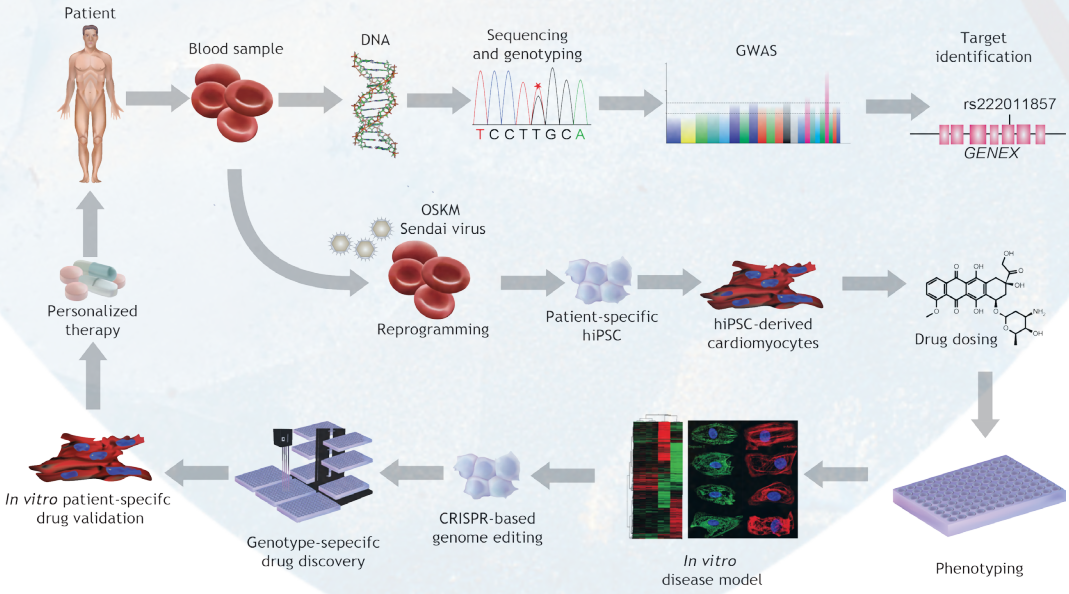
"The wealth of information furnished by a patient's genome makes genomic data a more valuable clinical tool than all other forms of patient information."

Large-scale academic GWAS studies in the pharmacogenomics field are commonly completed from 300-500 patients, using cost-effective genotyping arrays which assess ~700,000 SNPs and cost as little as $100 per participant. Although this strategy is useful, the SNPs identified are unlikely to be the true SNP of interest, merely one of the many SNPs co-inherited with it. As the costs of sequencing continue to decline, more of these studies will be completed with WGS, allowing researchers to close this data gap. WGS does have some drawbacks: analyzing 3 billion base-pairs means that each result is less statistically powerful than when measuring 700,000 base-pairs, so even more patients are required to complete a study.

Along with the promise of more readily available WGS, another major area of excitement is progress with the electronic health record (EHR). If all goes as planned, the EHR will be able to store patient data accurately and systematically, in a way that is compatible with the systems of different healthcare providers. Although progress with EHR implementation has been slow, the potential for researchers to recruit patients who have both comprehensive drug response EHR data and their WGSs available would allow GWAS to be performed virtually. This would populate databases with drug-genome correlations with improved statistical quality and eliminate the current substantial costs. The issue

"[We are] using the human induced pluripotent stem cell (hiPSC) model to take on this question of how to validate human SNPs of interest in drug response [...] looking at why some patients experience toxic damage to the heart as a result of their chemotherapy treatments."

with GWAS has always been validation and uncertainty that the correlations identified are real. GWAS findings are commonly not replicated in subsequent studies. Genetic questions have been successfully answered using model organisms, such as the mouse, but animal models are not suitable to experimentally validate human GWAS data, as the majority of GWAS 'hits' are not in protein coding regions but in the non-protein coding areas of the genome responsible for modulating gene expression. Non-protein coding areas are not necessarily under the same selective pressure to stay constant as those coding genes therefore have much more variation between species. In addition, the majority of model organisms are intentionally inbred, making them poor tools to study genetic variability. Therefore, for us to probe how these SNPs function, a human cell type is required, and for gene expression to be accurate this cell type needs to be as similar to the affected cell type as possible, for example if we want to study cardiac myosins, we need functioning cardiomyocytes. Finally, these cells must come from a patient with a specific phenotype, preventing the use of traditional cell lines such as HeLa cells.



My research group is using the human induced pluripotent stem cell (hiPSC) model to take on this question of how to validate human SNPs of interest in drug response. hiPSCs are made by taking a blood sample from a patient and temporarily inserting four reprogramming genes that cause the cells to switch from one cell type to another. The resulting hiPSC grow rapidly, are essentially immortal, and can theoretically turn (differentiate) into any cell type in the human body. Through years of careful optimization, we have refined our method of differentiating hiPSC to cell types affected by particular drugs, such as heart cells (cardiomyocytes) which contract or "beat" in cell culture, blood vessel lining cells (endothelial cells) that form tubes in culture, and neurons that can be electrically stimulated. Importantly, these cells maintain the genome of the patient from which they were generated and theoretically should also mimic that patient's drug response.

"We demonstrated that hiPSC-derived cardiomyocytes from patients who experience chemotherapy-induced cardiotoxicity are in fact more sensitive when treated with the same drug in the lab."

We have been looking at why some patients experience toxic damage to the heart as a result of their chemotherapy treatments. For doxorubicin, this cardiotoxicity occurs in ~8% of patients. The heart is a particularly interesting subject for hiPSC modeling because human heart cells are difficult to obtain from patients, as they are generally unwilling to undergo a clinically unnecessary heart biopsy, and heart cells are notoriously difficult to maintain in cell culture. We demonstrated that hiPSC-derived cardiomyocytes from patients who experience chemotherapy-induced cardiotoxicity are in fact more sensitive when treated with the same drug in the lab.

This finding confirms that patient drug sensitivity has a genetic basis because the genetic material is the only factor shared by the *in vitro* cardiomyocytes and the patients from whom the cardiomyocytes were derived. The hiPSC model also allows us to test whether SNPs previously found to be associated with certain effects are in fact causally related to the outcomes in question and to determine whether an effect was in fact caused by a target SNP or merely by a SNP co-inherited with the target. This research can be performed using genome editing tools to correct the SNP in question and see if the patient-specific drug response goes away. The most powerful current genome editing tool is called CRISPR/Cas9, this is a combination of an enzyme that cuts DNA (Cas9) and a guide to tell the Cas9 where to cut. Once the Cas9 has cut, we introduce a small template that the DNA is then tricked into using when repairing the cut, adding or removing our SNP of interest at the same time. We are currently working through massive amounts of GWAS data to experimentally validate previously discovered genome-drug effect relationships.

2018 © Dani Bergey
(UIC Biomedical Visualization graduate student)

"This finding confirms that patient drug sensitivity has a genetic basis because the genetic material is the only factor shared by the *in vitro* cardiomyocytes and the patients from whom the cardiomyocytes were derived."

Through hiPSC-validation research, it will one day be possible for a clinician to use genetic testing or existing WGS patient data to check whether a patient has genetic variants predicting that a particular drug will work especially well for the patient or will have potentially dangerous side-effects. Over time, the increasing availability of WGS data and EHR infrastructure will allow more SNPs to be discovered and experimentally validated, using hiPSC. A remaining question is whether clinical trials are required to further validate the drug-genome relationship or if the hiPSC-validation will be strong enough to stand alone. The potential elimination or reduction in clinical trials could allow drugs to be brought to market more quickly and reduce overall drug development costs with a substantial potential impact on patient care. Given sufficient availability of paired WGS and EHR data, this crucial validation process could be completed via efficient computational methods. Another issue will undoubtedly be that database companies may each interpret validated SNP data differently, as is currently the case with disease-related SNPs, making clinicians' responses dependent on their choices of vendor.

In summary, we now have powerful tools to answer pharmacogenomic questions in a thorough and actionable manner. Utilization of this approach will improve drug efficacy, reduce serial testing routine in clinical practice, and minimize unintended drug toxicity. Widespread adoption of precision medicine is not restricted by technological progress but by the uptake of genome sequencing and comprehensive digital health records as standard healthcare practice, both of which are necessary for researchers to vault us into the pharmacogenomic future.

Dr. Burridge is an assistant professor in the Department of Pharmacology at Northwestern University Feinberg School of Medicine. Dr. Burridge began his career in genomics and bioinformatics at the Sanger Institute working on the human and mouse genome projects. He completed a PhD in Human Stem Cell Biology at the University of Nottingham before pursuing postdoctoral fellowships at the Johns Hopkins University in Pediatric Oncology and then at Stanford University in Cardiology before becoming an Instructor in Cardiovascular Medicine at Stanford. For more than 15 years, Dr. Burridge has worked on the applications of human pluripotent stem cells (both hESC and hiPSC), concentrating on culture and differentiation methodologies, regenerative medicine, and disease modeling, specifically the pharmacogenomic and molecular mechanisms of chemotherapy-induced cardiomyopathy and heart failure.