



SimuGen Ltd: reliable, early prediction of drug toxicity with toxicogenomics, human cell culture and computational models

Quin Wills

SimuGen Ltd,
2 Oakington Business Park,
Cambridge, CB24 3DQ, UK
Tel.: +44 870 220 5910;
E-mail: quin.wills@
simugen.co.uk;
www.simugen.co.uk

Toxicogenomics is, arguably, the most exciting endeavor to better understand and/or predict the toxicity of drugs during their development, using technologies such as gene-expression microarrays. Through much of its (sometimes overzealous) build-up, toxicogenomics has found a natural niche as a bioinformatics and/or pattern-recognition tool, aiming to improve the mechanistic understanding of toxicity and animal-to-human extrapolation. The problem is that current approaches still need maturing and are expensive, slow, highly variable, often qualitative and do not easily yield information useful to decide whether to progress a drug candidate or not. Most crucially, they fail to address the main problem that industry faces – to be able to predict toxicity earlier in the discovery/development process. Rather than providing a conventional animal toxicogenomics service, SimuGen Ltd is launching a product that models gene expression in human cell culture, using an entirely novel approach to predict a remarkable spectrum of human *in vivo* (clinical) toxic end points, and to predict dose–toxicity relationships and molecular structure–toxicity relationships.

SimuGen is a computational biology company, developing software containing human toxicity models that accurately interpret molecular toxicology in human cell culture. The underlying novel algorithms extrapolate results to a broad spectrum of clinical (human) pathology, far beyond conventional cell culture predictive ability. SimuGen's approach is not conventional bioinformatics, as it does not output gene-expression results, but rather has novel 'under-the-hood' models to interpret the results. This exploits the characteristics of cell culture that are superior to animal toxicology (scalability and high result reproducibility) combining this with computational models that better align *in vitro* (cell culture) toxicity with *in vivo* outcomes. The analysis output (examples shown in Figure 1) is easily understood by nonspecialists and explicitly focused on facilitating the decision process, making for an ideal early-stage tool in industry to screen for toxic effects. SimuGen's current focus is liver toxicity, a leading problem stifling drug innovation; however, the basic approach will in future be applied to multiple *in vitro* models assessing multiple-organ toxicities.

The company was founded in 2006, drawing from the interdisciplinary talents out of the Cambridge (UK) genomics cluster. Funding commitments to date have been through private equity, venture capital and the UK Department of Trade and Industry. SimuGen's goal is to meet a major market gap that is not

adequately addressed by current toxicogenomics – to accurately predict toxicity in the early stages of drug discovery/development. The team strongly believes that current bioinformatics-and/or pattern recognition-based approaches suffer from major commercial drawbacks:

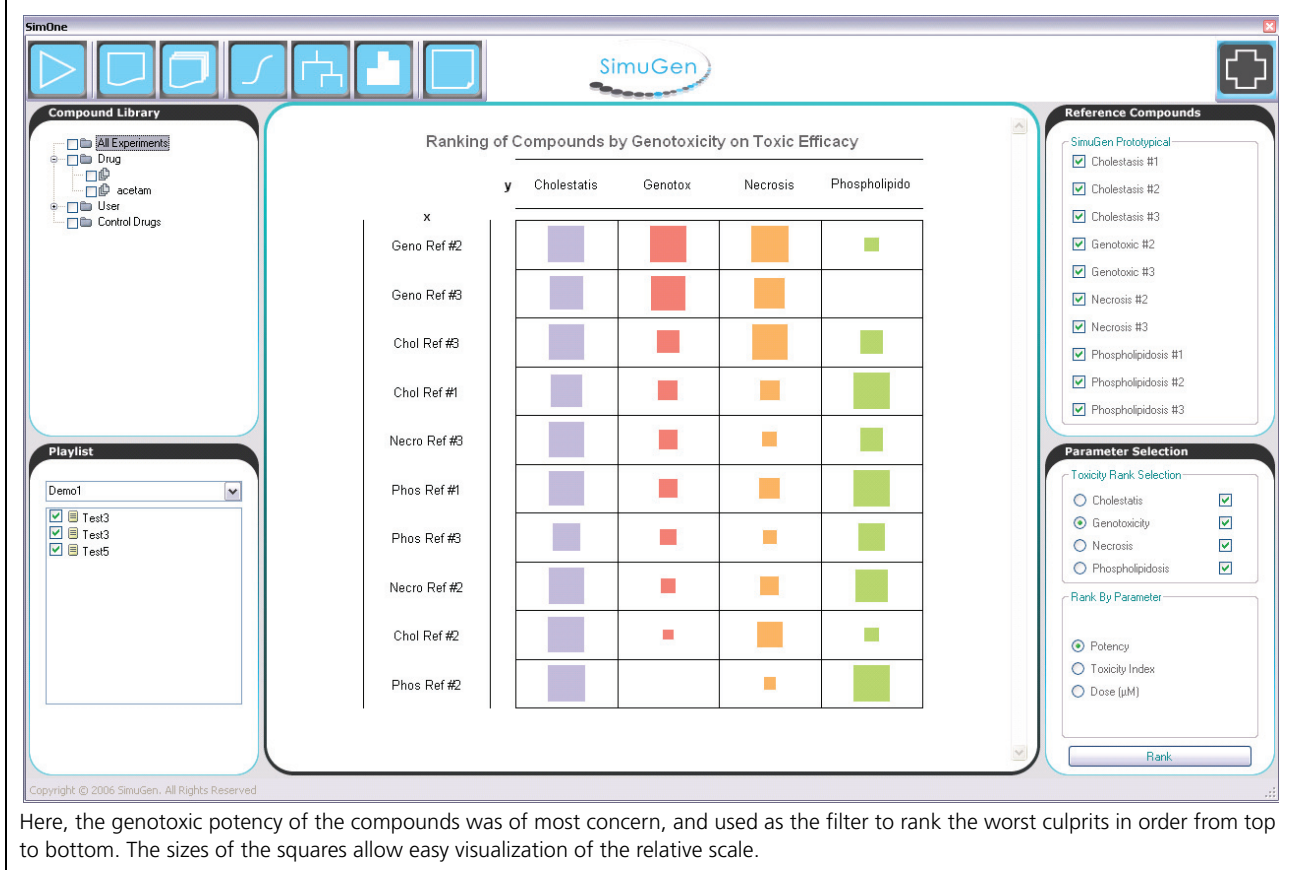
- Complicated and not typically focused on decision support
- Expensive
- Not scalable to higher throughput setups
- Not quantitative

SimuGen's novel approach addresses this and has been validated, with current work being the expansion of the models to predict a wider spectrum of clinical hepatotoxicities (liver toxicities). The flagship liver toxicity product will be launched in early 2008; the main company strategy being one of collaboration with contract research and toxicology services wanting to offer early-stage testing. SimuGen's philosophy is to focus on developing products around its core competency, applying novel and sophisticated computational biology models to select molecular biomarkers, growing its offerings and technology through codevelopment and marketing partnerships. The company maintains links with academic and industrial groups in toxicogenomics, *in vitro* toxicology and predictive absorption, distribution, metabolism, excretion and toxicity (ADMEtox). One such relationship is with LGC Ltd. The application of

Keywords: gene-expression profiling, *in vitro*, liver toxicity, predictive toxicology, toxicogenomics

future
medicine part of fsg

Figure 1. The intuitive ranking of candidate drugs in a chemical series according to various toxicities of interest.



Here, the genotoxic potency of the compounds was of most concern, and used as the filter to rank the worst culprits in order from top to bottom. The sizes of the squares allow easy visualization of the relative scale.

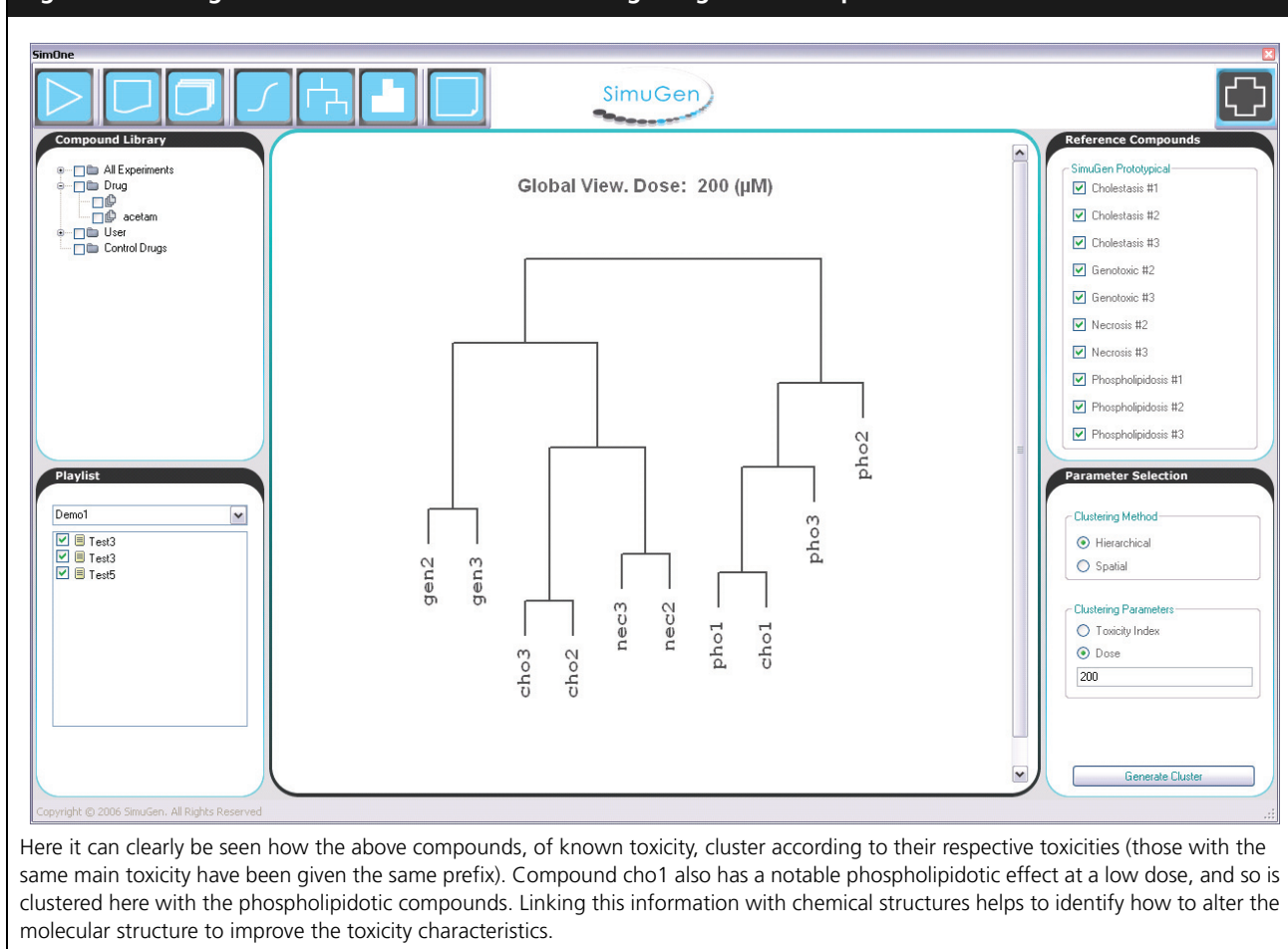
their in-house cell culture, molecular-profiling research and technology skills, supported by their national position as the UK National Measurement Institute in this field, ensures the validity and robustness of SimuGen's software modeling and specific address of the market gap.

Toxicity & toxicogenomics

Next to drug efficacy, drug toxicity has been highlighted as the main reason for candidate molecule failure in drug development. The crucial need to focus on toxicity was highlighted in a, now often-cited, US FDA white paper [101]. The burden on other industries has also been brought to the fore with the much publicized new EU Registration, Evaluation, Authorization and Restriction of Chemicals (REACH) legislation [102]. REACH has been argued to be one of the most important EU legislations to date, requiring the testing of tens of thousands of chemicals, with expected global knock-on effects. Toxicogenomics has been seen as a 'better, cheaper and faster' tool

to greatly improve current toxicology, speed up chemical/drug development and reduce the burden on animal testing. As with many 'omic' technologies, there has been much hype, but it has begun maturing primarily as two types of gene-expression applications. The first is microarray bioinformatics, where gene expression can be used to identify (often poorly understood) toxic processes, or to better discover underlying mechanisms. The second is as a database-driven pattern-recognition tool. Gene-expression profiles might be seen as unique drug 'fingerprints', the idea being that if a fingerprint closely matches that of similar toxicants in a database, there is suggestion of toxicity. However, this is only 'toxicity by association', and how novel and robust enough these fingerprints are for useful predictive tools in any large-scale way has yet to be seen. Both of these approaches tend to be expensive, slow, complicated and fairly qualitative. As the science matures, they hold promise in mechanistic toxicology and improving animal-human extrapolation.

Figure 2. Dendograms are one method of visualizing the global toxic pattern in a chemical series.



Here it can clearly be seen how the above compounds, of known toxicity, cluster according to their respective toxicities (those with the same main toxicity have been given the same prefix). Compound cho1 also has a notable phospholipidotic effect at a low dose, and so is clustered here with the phospholipidotic compounds. Linking this information with chemical structures helps to identify how to alter the molecular structure to improve the toxicity characteristics.

SimuGen's approach & benefits

SimuGen's aim is not to improve mechanistic understanding through bioinformatics, or to provide a database-driven tool. The company addresses the need for informative early-stage toxicology through applying toxicogenomics to human cell cultures in a novel way. The end user cultures a molecule of interest, together with cells, for 24 h over multiple concentrations, the software input being quantitative PCR values of select genes. This means a predictive toxicology tool that can be run in a standard cell and molecular laboratory, only requiring small amounts of the tested compound (versus animal testing), and which is rapid and robust.

The core technical novelty is SimuGen's combination of genes in proprietary toxicity models, which monitor how the genes collectively change across drug concentration, rather than the more standard utilization of raw expression-data 'snapshots' as rudimentary biomarkers. The parameters from these models can then be further

input into pattern-recognition approaches (as opposed to just raw expression data) to discover biologically meaningful patterns. To the end user, this provides three important pieces of information, described below.

Predicted spectrum of human toxicity

SimuGen's impressive augmentation of cell-culture results beyond simple necrosis has been validated against compounds of known toxicity, and externally reviewed. The full product for launch in early 2008, will provide information on ten human liver toxicities from acute toxicity to more liver-specific toxicities (e.g., cholestasis) and longer-term effects (e.g., genotoxic carcinogenesis).

Dose-response relationship

It is not sufficient that applications only suggest likely toxicities. Dose makes the toxin, and it is crucial to predict the dose-response relationship. Compounds of interest in a chemical series can then be ranked on toxic potency, and the best candidates taken forward. Figure 1 demonstrates a

screenshot of SimuGen's current software, with a rapid-ranking feature for compounds over four toxicities of interest.

Structure–toxicity relationship

Pattern-recognition approaches can be applied to a chemical series using parameters (such as potency) from the models to provide more biologically meaningful 'global views' of the toxic patterns across all the compounds being tested. Tied in with the molecular structures, this is an essential tool for medicinal chemists, suggesting what could be done to the molecules to result in more acceptable toxic parameters. As a result, rather than providing a post-mortem of results, this approach better helps those higher up the pipeline make fewer toxic compounds. Figure 2 demonstrates a simple dendrogram of the compounds tested in Figure 1.

Future perspective

SimuGen's strong interdisciplinary expertise, ranging from clinical medicine to computational biology, has proven essential in the current pioneering development of a predictive liver-toxicology tool. The company vision is to continue expanding partnerships along three paths:

- Integration into early stage absorption, distribution, metabolism, and excretion (ADME)
- Application to different cell-culture models for further target-organ toxicities
- Providing the current pharma-focused tools to other industries requiring early-stage toxicity testing

Financial disclosure

QW is employed by SimuGen Ltd, Cambridge, UK.

Highlights

- Current toxicogenomics approaches to predict drug-candidate toxicity are still maturing, are slow, expensive, highly variable and often qualitative, do not explicitly guide the decision to progress good candidates, and (most crucially) fail to allow toxicity prediction early in the discovery/development process.
- SimuGen Ltd is a highly interdisciplinary company founded in 2006, based in the well-known Cambridge (UK) genomics cluster, with the goal of specifically addressing these shortcomings.
- Funding commitments to date have been through private equity, venture capital and the UK Department of Trade and Industry, with the flagship product launch (liver toxicity) being in early 2008.
- SimuGen's novel, validated approach is not standard animal toxicology bioinformatics; it models human cell-culture toxicity to accurately predict a remarkable spectrum of organ-specific human toxicity.
- SimuGen's software is aimed specifically at providing intuitive output to guide early-stage toxicity screening and molecule synthesis by predicting dose–toxicity and structure–toxicity relationships.
- SimuGen's philosophy is to focus on its core computational biology competency, expanding current development and marketing partnerships to increase absorption, distribution, metabolism and excretion (ADME) market penetration, increase the spectrum of predicted target organ toxicities and expand its pharma-focused tools into other industries requiring early-stage toxicity testing.

Websites

101. Innovation or stagNation: CHALLENGE and opportunity on the Critical Path to New Medical Products www.fda.gov/oc/initiatives/criticalpath/whitepaper.html

102. Environment – Chemicals – Registration, evaluation, authorization and restriction of chemicals (REACH) http://ec.europa.eu/environment/chemicals/reach/reach_intro.htm