Proteomics, Informatics, & Data Mining to Reduce Costs of Drug Development for Substance Use Disorders

A shared modality of disease—whether caused by infectious or chemical agent—is phenotype. Phenotype is a *material read-out* with a distinct advantage concerning disease: biological units mediate pathologic processes. As such, biomolecules with function in disease processes are targets for therapeutic intervention. Some of these targets and concomitant potential therapeutics lie unutilized within current molecular data. At Viralchemy Bioscience (VBS), we have engineered and implemented our protein-centric informatics technology aggregating presently available biological data to identify high-priority targets and therapeutic candidates. By automating and synergizing available libraries and databases, we can perform our automated discovery (AuDisco) data mining strategy on any known protein and >90 million molecules for pharmaceutical discovery efforts. While our foundational studies have been directed towards the discovery of broad-spectrum antiviral agents (BSAA), our versatile technology has the capability to be easily redeployed to provide groundbreaking insight to the neurobiology of SUD, an area in dire need of targeted therapeutics and individualized medicine.

The application of 'omics' discovery tools is now ubiquitous in disease research, including the neurobiology of SUDs. Dynamics in the molecular neurobiology resulting from the administration of addictive substances is most robustly characterized at the genomic and proteomic level. Multiple studies with high quality data primed for analysis with our technology exist. A few examples include human cocaine use[1], rodent amphetamine self administration[2], and morphine dependence[3]. In human cocaine use, 11 nucleus accumbens cytosolic proteins were positively identified and up-regulated in cocaine overdose verse control patients—a condition with neuropathology similar to Parkinson Disease[1]—representing potential therapeutic targets with immediate neurobiological importance. In a rodent model of naïve, binge, voluntary abstinence, and relapse of amphetamine usage, 1,294 differential proteins were detected and 73 positively identified hippocampal proteins were shown as significantly different across the usage states[2]. Also, in a rodent model of morphine usage, 19 proteins were positively up-regulated in correlation with morphine-dependence[3]. Aggregating hundreds to thousands of these proteins with roles in the neurobiology of SUD will present novel targets for small molecule treatment of substance abuse, as seen in cocaine use[4] and acute treatment with Naloxone for opioid overdose. These studies offer examples of the data upon which we construct our architecture, using available 'big data' and leveraging modern computational power to mine for small molecules with therapeutic potential.

While our technology handles the mechanics of our discovery pursuits, the collective and multidisciplinary experience of *our team encompasses a startup taking a diverse view of a singular mission: make drug discovery more efficient.* As it now consumes on average $2.5 billion and 10 years to bring a new drug to market[5], a practical method to save both valuable capital and lives is intelligent, resource-efficient discovery in preclinical pharmaceutical discovery. Sadly, the victims of inefficiency in the current drug development processes hindering treatments from coming online are patients; the 10-year delay of potentially live-saving therapeutics and the cost burden of treatments is increasingly placed on them. We argue VBS technology can address this need by guiding the construction of small molecule libraries underutilized for potential applications in the treatment of SUD with characterized in silico hits. With 23.5 million patients annually needing treatment for illicit drug/alcohol use[6] and the diversifying known and unknown chemicals and analogues exploited as recreational drugs (e.g. carfentanil), identification of molecular targets conserved across various illicit substance uses has the potential to reveal broad-spectrum targets for the neurobiological treatment of SUD. Small molecule treatment of SUDs is showing increasing promise and garnering substantial clinical and financial interest[7]. Our versatile technology is applicable to multiple drug discovery programs for many diseases. *Thus, Viralchemy Bioscience has the capability to enter numerous markets ranging from viral diseases to SUD, further giving our team the potential to become a*

successful venture that will not only disrupt the status quo but enhance drug discovery and as a result, patient treatment.

Our interdisciplinary and highly experienced team is composed of: Trevor Gale, Ph.D., Tim Horton, M.S., and Ben Bradley, Ph.D. (MD 2016). Trevor, a Systems Biologist, has over 6 years experience in using analytical chemistry to discover prognostic small molecule biomarkers from clinical matrices. Trevor's capabilities integrate multiple disciplines from chromatographic and mass spectrometric analysis to molecular biological manipulations for the discovery of small molecules with novel properties. Tim, a Senior Software Engineer, has over 15 years of experience in statistical and data processing languages and informatics tool development for novel biological motif identification. Tim's extensive data science capabilities reduce the noise of big data to help deliver results from convoluted systems. Ben, a Physician Scientist, utilizes diverse molecular biology techniques to characterize targetable disease phenotypes. Ben's capabilities include translating clinical need to fundamental research and vice versa. Ben's focus has included molecular interactions of neutralizing antibody epitopes, efforts ultimately directed towards immunotherapeutics development. *As a team, we operate at the intersection of biomedical and technical sciences, integrating unique yet complimentary fields to pioneer a novel tool to improve pharmaceutical development efficiency.*

The cornerstone of our technology platform is synergy of big data for a specific disease. The advantage of using proteomics data is *material readout,* i.e. analyzing proteomes with our technology produces a target priority list of immediate therapeutic relevance. Our algorithm integrates the VBS target priority list and generates in silico therapeutic candidates identified from databases containing over 90 million small molecules with characterized bioactivity results. As stated previously, *the platform can easily be repopulated with data pertinent to any specific disease—including SUD—and redeployed to produce a proprietary target list and promising but underutilized drug candidates.*

In our technology, the build stage is aggregation of big data ('-omics') and generation of a disease specific architecture. The build process is a combination of automated and manually curated data identification. From this architecture, multiple user interfaces are available to toggle between disease and target. For SUD, this would allow one to compare profiles between distinct addictive substances, examine global trends, or list targets by summation for the entire dataset or by discreet substance [condition]. Our user interfaces that will be tailored to SUD are shown in a prototype built on viral disease proteomics data (**Figure 1, left**). From this architecture, the Viralchemy algorithm initiates AuDisco, an automated query of the highest priority targets against >90 million small molecules (**Figure 1, right**). AuDisco specifically identifies and prioritizes hits with activity against these targets. The AuDisco strategy also integrates available intellectual property (IP) information. *By bundling preexisting activity, toxicity, and IP data, AuDisco streamlines the likelihood of identifying small molecules with high therapeutic and market potential, fulfilling a drive towards both novel and repurposing drug discovery*.

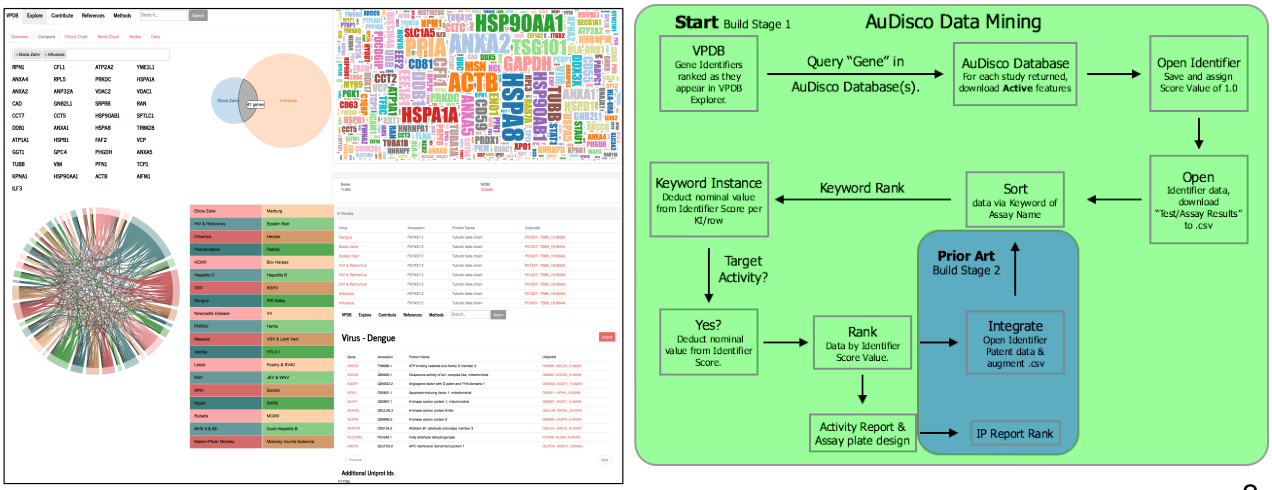

Figure 1. Viralchemy User interfaces (left) and Automated Discovery algorithm (right).

Our team has produced evidence of a working prototype through a proof-of-concept study with the architecture built on proteomics data of extracellular viral particles. The target list, presented as NCBI Gene Identifiers and generated from >2,000 features, prioritizes host proteins implicit in viral infection. The AuDisco algorithm then queries the identifiers against sources containing relevant biological and patent data. Discreet results are scored through the AuDisco algorithm and returned in a spreadsheet format with an option to assemble multiwell plates with available small molecules (>100 million).

As a proof-of-concept for the discovery of small molecules with antiviral activity, 5 VBS hits were selected and screened against a Human Immunodeficiency Virus (hereafter pseuodparticle, PP). The rationale for this experimental design is that AuDisco identified inhibitors modulating the *targeted host protein* result in lower signal (less infectious or less total virus) compared to uninhibited control. Targeted proteins can eventually act as validated targets and, in the context of viral infection, offer the distinct advantage of lower probability for the generation of resistance. Of the 5 compounds tested, 2 compounds were antiviral; a compound denoted VRL5 was the most effective. In the beta version of the prototype, this cellular enzyme is registered in 8 unique instances, highlighting the potential conservation of host processes across diverse disease-causing agents. VRL5 exhibits a dose-dependent antiviral activity in the PP screen (**Figure 2**). The only observed reduction in cellular viability was at the highest concentration tested (48% viability at 100 μM) and only in 1 of 3 cell types, i.e. completely non-toxic in 2 of 3. Selectivity indices for VRL5, in this context, specific antiviral activity, are within therapeutic range (**Table 1**).

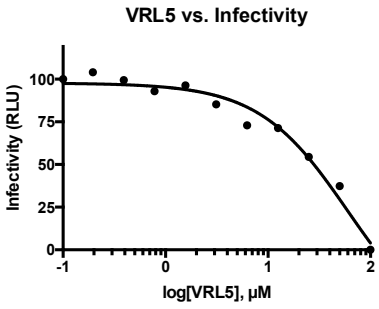

Figure 2. A Inhibition of PP system.

Table 1	Cellular CC_{50} (μM)	PP IC_{50} (μM)	Selectivity Index (CC_{50}/IC_{50})
HEK293/T-17	221.8	60.2	3.68
MDCK	2268		37.67
VERO-E6	151.9		2.5

In competent viral systems, VRL5 *significantly and reproducibly* inhibited Dengue virus plaque reduction at concentrations between 12.5 and 50 μM. In plaque forming unit (PFU) reduction assay—a measure of inhibitor activity on viral progeny—VRL5 dose dependently and significantly reduced Dengue virus between 25 and 50 μM. At 100 μM, VRL5 completely inhibited Zika virus plaque formation and had no observable toxicity (**Figure 3A**), consistent with the lack of toxicity observed in the assay cell type (at 50 μM 109% viability). An inhibitory concentration between 50 and 100 μM is consistent with nonlinear regression analysis of PP inhibition returning a IC_{50} of 60.2 μM. In PFU-reduction assay at all concentrations tested, VRL5 showed significant reduction of Zika virus progeny with p = 0.003 at 50 μM (**Figure 3B**). *Our proof-of-concept results illustrate VBS discovery: agile identification of novel targets and high priority therapeutic candidates.*

Additional evidence of the utility of VBS discovery is illustrated by the fact that VRL5 has pre-existing in vivo data. Administration of this molecule is safe in rodents, non-human primates, and humans, having been safely dosed in Phase I/II clinical trials for non-infectious genetic and neurodegenerative conditions (references withheld as they would identify VRL5). Therefore, we argue *our technology can be a highly complementary component to pharmaceutical discovery efforts by identifying high priority targets and candidate drugs early in the discovery process. This technology can serve to identify novel small molecule treatments and foster an effort of drug repurposing, leveraging unutilized therapeutic potential from drugs with current FDA approval and potentially mitigating the front end investment necessary to move a drug lead down the development pipeline.*

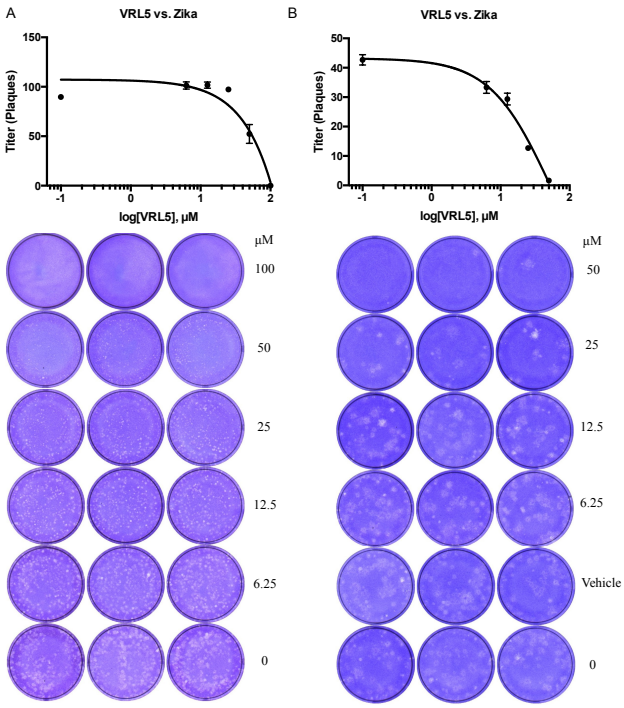

Figure 3. A Plaque reduction assay and **B** Plaque forming unit reduction assay with Zika virus vs VRL5.

A new startup should map a strategy to incrementally chip away at the available customer base; our entry strategy is *search and deploy*. We have identified small-to-mid level preclinical and early clinical stage pharmaceutical companies developing small molecule therapeutics as our target market. As a first approximation of a defined customer profile, we will execute a focused outreach effort, beginning with pre-screened participants at a select number of companies (<5). We will run VBS discovery for validated (non-proprietary) targets provided by the company and return a blinded report summarizing our findings. In exchange for participation/input—if desirable—we will share our findings. E.g. there are multiple biopharmaceutical startups focused on the human bodies' interaction with the microbiome. We believe that by identifying host proteins dynamic as a result of microbiome dysregulation, our technology can identify drug candidates not currently recognized for potential use in these disorders. By working with companies, we can define deliverable services and develop an educated customer profile. By executing this lean model strategy we will apply our technology and receive real-world, real-time feedback on the utility our technology provides which will allow us to gauge the viability of our company as it fits within the existing pharmaceutical infrastructure.

Given the exorbitant expenditures of pharmaceutical development, our technology offers a cost mitigation strategy for our customer at the front end of the pharmaceutical development spectrum. This can be a complementary technology to ensure the investment only in optimized preclinical candidates prior to progression down the development pipeline, i.e. consuming resources in the cash flow "valley of death." Since real dollars cannot be saved in the clinical trial phases, it is exceedingly important to reduce spending in preclinical investigation.

Finally, VBS will endeavor to define and refine our proof-of-concept studies. The identification of VRL5, a small molecule with potent, broad-spectrum antiviral activity that is extremely well tolerated in humans, offers a stellar example of the potential of our technology as a consummate drug repurposing strategy. Leveraging computational power and our software, *Viralchemy can identify druggable molecules possessing background data that may completely allay the need for substantial preclinical testing for a diverse array of diseases.*

References

1 Hemby, S. E. Assessment of genome and proteome profiles in cocaine abuse. *Prog Brain Res* **158**, 173-195, doi:10.1016/S0079-6123(06)58009-4 (2006).
2 Freeman, W. M. *et al.* Distinct proteomic profiles of amphetamine self-administration transitional states. *Pharmacogenomics J* **5**, 203-214, doi:10.1038/sj.tpj.6500309 (2005).
3 Kim, S. Y. *et al.* Proteomic analysis of phosphotyrosyl proteins in morphine-dependent rat brains. *Brain Res Mol Brain Res* **133**, 58-70, doi:10.1016/j.molbrainres.2004.09.018 (2005).
4 Schmeichel, B. E., Herman, M. A., Roberto, M. & Koob, G. F. Hypocretin Neurotransmission within the Central Amygdala Mediates Escalated Cocaine Self-Administration and Stress-induced Reinstatement in Rats. *Biological Psychiatry* (2016).
5 DiMasi, J. A., Grabowski, H. G. & Hansen, R. W. Innovation in the pharmaceutical industry: New estimates of R&D costs. *J Health Econ* **47**, 20-33, doi:10.1016/j.jhealeco.2016.01.012 (2016).
6 Abuse, N. I. o. D. *Drug Facts: Treatment Statistics*, (2011).
7 Ballard, J. Shreveport Pharmaceutical Company Receives $11 Million Grant. *Silicon Bayou News* (2016).