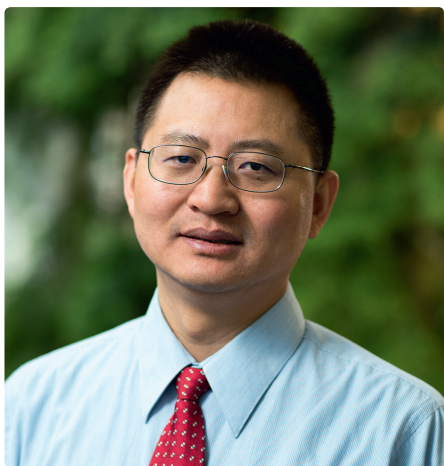


Analysing multidisciplinary vaccine mechanisms

Dr Yongqun He describes the importance of informatics in his work on vaccine research and design, including the benefits of a multidisciplinary background and the team's future research activities



To begin, what are the main objectives of your investigations? What specific biomedical questions are you seeking to address through your research?

The primary aims of my research are to develop and apply bioinformatics and systems biology approaches to address fundamental biomedical questions, especially in the areas of immunology, vaccinology, ontology and biological pathway analysis.

Alongside your appointment in the Department of Microbiology and Immunology, you also have a position at the Unit for Laboratory Animal Medicine (ULAM). How do you balance these two interests?

My primary position is at ULAM, which does not require me to conduct clinical services. Instead, I can fully focus on research, which I find beneficial. I also have appointments in the Center for Computational Medicine and Bioinformatics as well as the Comprehensive Cancer Center. My research centres on microbiology, immunology and bioinformatics, with vaccinology forming the primary focus of my laboratory research. The bioinformatics research includes vaccine informatics, biomedical ontology development and applications, literature mining, and Bayesian network analysis of biological pathways. My wet lab research focuses on the study of basic mechanisms of *Brucella* vaccines, and translational *Brucella* vaccine development.

My strategy of balancing different interests is to integrate different experiences and expertise to perform interdisciplinary

research. This is a unique angle, and novel in terms of methodology.

Could you offer insights on your work to improve vaccine informatics for the future?

With explosive big datasets, data integration and automated reasoning are crucial. A major foundation of computer-assisted data integration and reasoning is the establishment of community-based ontologies that logically represent the concepts and relations in specific biomedical domains. We have founded and led the development of several biomedical ontologies including the Vaccine Ontology (VO) and the Ontology of Adverse Events (OAE). These ontologies have laid out a foundation for generating a vaccine linked data system. Novel data analysis tools, for example, ontology-based literature mining and Bayesian network modeling, can further be developed. Such a strategy will support advanced host-vaccine/pathogen interaction analysis, fundamental biological pathway exploration, and rational vaccine design. This is a project that we would like to develop in the future.

What does the term 'reverse vaccinology' (RV) denote? How does your approach lend itself to this?

RV is a revolutionary new vaccine development strategy that starts with bioinformatics prediction of vaccine candidates. A milestone in the RV strategy is that, in just over 10 years, the research by Dr Rino Rappuoli and his group has resulted in the first RV-derived meningitis vaccine: Novartis's Bexsero. While the RV strategy has been proven successful in many vaccine development studies, it is time consuming and inefficient to predict vaccine candidates without a systematic bioinformatics pipeline.

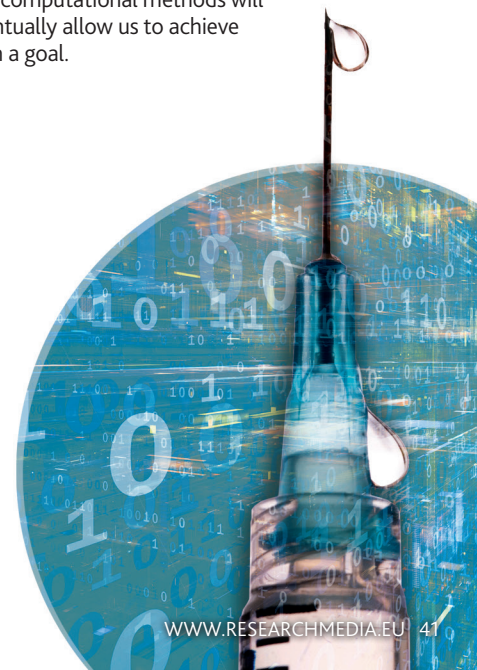
We developed Vaxign, the first web-based vaccine design program, based on genomic sequences utilising the RV strategy. Predicted features in the Vaxign pipeline include protein subcellular location, transmembrane helices, adhesin probability, sequence conservation among pathogenic strains, sequence exclusion from genome(s) of nonpathogenic strain(s) and epitope binding to MHC class I and class II. Vaxign has been demonstrated to successfully predict vaccine candidates for *Brucella spp*, uropathogenic *E. coli*, herpes simplex viruses, and several other pathogens.

One area of translational computational immunology research that is progressing rapidly to the clinic is cancer vaccine design. What activities are you undertaking in this field and what results have been produced?

Our VIOLIN database has started to collect cancer vaccine data. Systematic analysis of this data will allow us to find new consistent patterns useful for rational cancer vaccine design. It is also possible to update our Vaxign program to include a new function of rational cancer vaccine design. Furthermore, we are in the process of analysing gene expression transcriptomics and proteomics data for more efficient vaccine design, including cancer.

Do you have any final thoughts regarding the state of the art of bioinformatics?

I believe that a life is fundamentally one single network that is composed of genes, proteins, and metabolites, and how they interact with each other and with the environment of the living organism. Different genetic factors may change the behaviors and phenotypes of a living organism to environmental factors. It is challenging to integrate all existing data to uncover the one network and how it interacts with environmental factors. However, it will be rewarding to build up such one network using existing data and enable the prediction of new pieces of the network. Ontology-based data integration and advanced statistical and computational methods will eventually allow us to achieve such a goal.



Immunology and informatics

Researchers from the **University of Michigan** are focusing on using modern bioinformatics to improve the efficiency of vaccine research and design. Such an approach holds promise for immunologists worldwide and is enabling the lab to drive forward their own research into a human vaccine against *Brucella*

VACCINES HAVE SAVED more lives across the world than any other health intervention, with the exception of clean water. Since Edward Jenner, an English country doctor, produced the world's first vaccine against smallpox, the race has been underway to produce vaccines to prevent as many diseases as possible.

Modern vaccinology has advanced notably since the time of Jenner. The production of vaccines is now based as much on genetic analysis and computational prediction as it is trial-and-error. Long gone are the days of the early pioneers like Louis Pasteur and Henry Touissant, who tested prototype vaccines on patients with little or no knowledge of the potential outcomes. Today's vaccine scientists enjoy the benefits of modern techniques and tools employed to design effective vaccines, both making production safer and more efficient.

One such scientist working to further improve the resources and tools available to vaccine researchers is Dr Yongqun He from the University of Michigan's Medical School. He arrived at the field of immunology

after attaining qualifications in veterinary, immunology and computer sciences, placing him in a strong position to further this multidisciplinary area of research. This background has enabled him to identify and design effective computational and predictive tools for research into potential vaccines. He and his colleagues are concerned with two primary research objectives: the creation and improvement of vaccine informatics tools, and the use of these tools to study vaccine mechanisms and produce their own vaccine for the prevention of *Brucella* infection.

REVERSE VACCINOLOGY

The lab has approached the production of new vaccines using a modern strategy called reverse vaccinology (RV). Traditional methods of vaccine production rely on testing weakened (i.e. attenuated) pathogens, or cleaving antigenic particles from a virus or bacteria for testing. This process involves much trial and error, necessitating huge investments of time and money in order to produce a safe and usable vaccine.

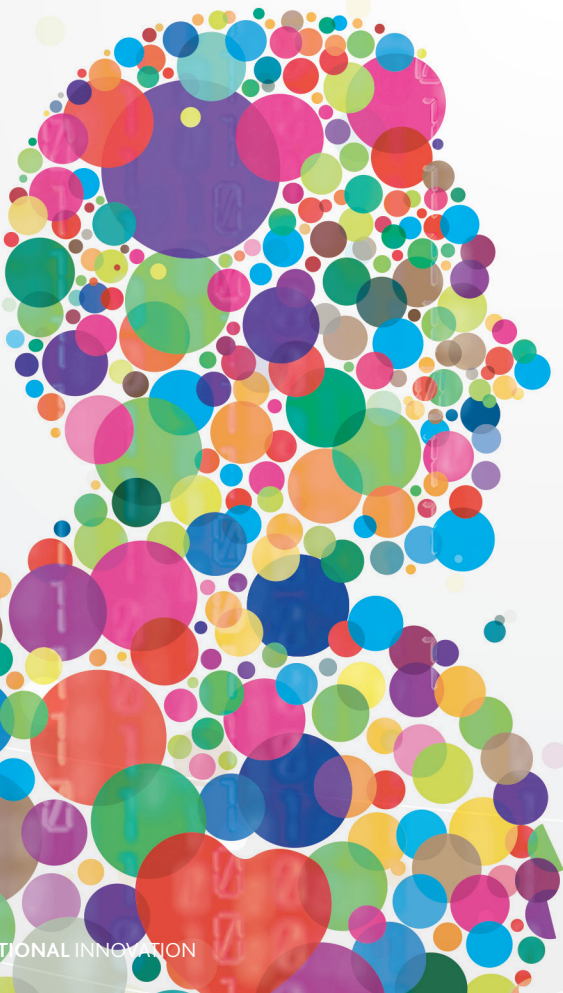
RV is able to overcome several of these shortfalls by predicting likely antigenic particles based on genomic analysis of the pathogen in question, as He elaborates: "Instead of starting in the wet lab, a more efficient way to identify vaccine candidate proteins/genes is to begin with the analysis of genome sequences and prediction of vaccine candidates, followed by experimental verification". Far from being a theory, RV is in use today and has produced useable human and animal vaccines. The approach was first employed in the production of a vaccine for serogroup B *Neisseria meningitides*, resulting in the existence of an RV-based product.

The world of vaccinology is becoming increasingly multidisciplinary, as evidenced by He's approach. The Michigan team, a group from Texas A&M University and a biotechnology company, have formed a collaboration, designed to enable the production of a recombinant vaccine against African Swine Fever Virus (ASFV). The project, funded by the US Department of Homeland Security Science and Technology, has enjoyed notable success. The researchers were able to use a bioinformatics approach to identify and analyse potential vaccine targets: "We used our algorithm to rank all the possible virus proteins as potential vaccine targets. After discussion with the collaborators, a subset of protein targets were down selected for experimental vaccine development," highlights He. Many of these identified proteins have now been purified in a laboratory setting and shown to be reactive with ASFV antiserum. Evaluation of these recombinant proteins in animal studies is ongoing.

THE IMPORTANCE OF INFORMATICS

The central theme of the Michigan group's work is the use of modern informatics and technology to enable and streamline both their research and that of the wider vaccinology profession. He's work in this area has been prolific. Highlights include the production of the first

The central theme of the Michigan group's work is the use of modern informatics and technology to enable and streamline both their research and that of the wider vaccinology profession



He lab members

Front row:

Abra Guo, Asiyah Yu Lin, Erica Marcos, Rebecca Racz, Shelley Zhang, Sirarat Sarntivijai
Back row: Omar Tibi, Oliver He, Joseph Ostrow, Gerry Hish, and Desikan Jagannathan.



Previous lab members (not in the photo) include many graduate and undergraduate students, DVM summer interns, postdoctoral research fellows, visiting scholars, and Zuoshuang (Allen) Xiang, a major bioinformatician who worked in the He lab for eight years.

research-orientated comprehensive vaccine database and analysis system, VIOLIN, along with the creation of web-based programmes for the use of immunology research teams across the world. Perhaps the most versatile of these is Vaxign – web-based software which enables users to explore the potential of RV methods in their own research.

One of the most important but often undervalued jobs in the production and monitoring of new vaccines is pharmacovigilance – this term refers to the observation of vaccine safety and the recording of any side effects or adverse events when a vaccine is in use. Understanding the data collected in the general population after a product is released is essential and allows researchers to build a better picture of the risks or absence of risks for each vaccine. To enable such work, He and colleagues have also developed the Ontology of Adverse Events (OAE) as a data infrastructure to aid bioinformatics applications. OAE can also be used to study drug pharmacovigilance.

THE BURDEN OF BRUCELLA

Brucella is a parasitic bacterium which can replicate inside cells, making it difficult to treat with antibiotics. Infection is most commonly acquired from unpasteurised milk or cheese. Most patients fully recover with timely but prolonged treatment. However, without early diagnosis and therapy the disease brucellosis can often lead to chronic infections, and sometimes infected individuals will die. *Brucella* has been a research focus for over 70 years, and in 1954 the bacteria became one of the world's earliest biological weapons following research and manipulation in the US. Today, He and colleagues are attempting to produce an effective vaccine against infection in humans. He explains the focus on this bacterium: "*Brucella* is one of the most

common zoonotic diseases in the world in humans and a variety of animal species".

Wild-type *Brucella* is known to infect macrophages – altering their behaviour and preventing cell death. In comparison, attenuated cattle *Brucella* vaccine strain RB51 and many other rough attenuated strains trigger a novel caspase-2-mediated pro-inflammatory cell death (pyroptosis) of infected macrophages, making them an easier target for immune responses. He and the Michigan team first discovered and are actively studying this cell pathway: "Our research aims to elucidate the caspase-2-mediated pro-inflammatory cell death pathway and its biological effect on microbial pathogenesis, protective immunity against brucellosis and other diseases". Ultimately, this will inform and enable the team's production of a vaccine against this widespread burden.

GOALS FOR THE FUTURE

He is now looking at ways of integrating and analysing data relevant to the field of immunology: "One way to improve vaccine informatics is to improve vaccine data integration through ontology-based Linked Data (or the Semantic Web). Based on the integrated data, more systematic and advanced computational methods can be developed to improve vaccine mechanism study and rational vaccine design".

The work of He's lab represents considerable progress towards the future prevention of *Brucella* infection. Perhaps more importantly, it also provides us with insight into the possible future of immunology and vaccine production. If He's bioinformatics approach can provide ergonomic tools which improve the efficiency of vaccine design and critically reduce the cost, we may soon enjoy the benefits of a whole range of new human vaccines for currently unchecked pathogens.

INTELLIGENCE

VACCINE INFORMATICS

OBJECTIVES

- Analyse fundamental host-vaccine/ pathogen interaction mechanisms and support rational vaccine design using bioinformatics and systems biology approaches. *Brucella* is He's wet lab bacterial research model
- Develop and apply biomedical ontologies and ontology-based informatics tools to support data integration, literature mining, and pharmacovigilance analysis
- Biological pathway analysis using ontologies and Bayesian network algorithms

PARTNERS

Darrell Abernethy, Brian Athey, Mangkey Bounpheng, Marcelline Harris, Alfred Hero, Gary Huffnagle, Harry Mobley, Jean Nemzek, Mary O'Riordan, Arzucan Ozgur, Helen Parkinson, Steve Qin, Barry Smith, Cui Tao, Peter Woolf and Qingmin Wu, among others

FUNDING

National Institute of Health – National Institute of Allergy and Infectious Diseases (NIAID)

US Department of Homeland Security Science and Technology

CONTACT

Dr Yongqun He (Oliver)
Principal Investigator

University of Michigan Medical School
Unit for Lab Animal Medicine, Department of Microbiology and Immunology, and Center for Computational Medicine and Bioinformatics
018 ARF, 1150 W Medical Center Dr
Ann Arbor, Michigan 48109
USA

T +1 734 615 8231
E yongqunh@med.umich.edu

YONGQUN HE (OLIVER) received his PhD in Infectious Diseases & Immunology from Virginia Polytechnic Institute and State in 2000. He has been Associate Professor at the University of Michigan Medical School since 2011.

