

## WUXI EX MACHINA

BY WINNIE PONG, STAFF WRITER

Since its acquisition two years ago by **New WuXi Life Science Ltd.**, **WuXi NextCode Genomics Inc.** has been busy building out its artificial intelligence (AI) capabilities and test driving them via its expanding network of academic and industry collaborators. The strategy is starting to pay off with both a publication and presentation in the last two months showcasing the platform's diagnostic power and its ability to accelerate discovery of new mechanisms in cancer.

WuXi NextCode has already launched three genomics-based diagnostics companies in China, and last month raised a \$75 million series B round, which will fund commercialization of the tests, as well as continued development of its deepCODE AI tool.

On May 3, the company published a study in *Nature* in partnership with a **Yale University** group using the AI platform to trace the role of FGF signaling in vascular development to effects on glycolysis, a cellular function the FGF pathway had only been tenuously linked to before.

And at April's meeting of the **American Association for Cancer Research (AACR)**, WuXi NextCode showed deepCODE could classify 9,000 tumors into 27 of The Cancer Genome Atlas (TCGA) tumor subtypes with over 95% accuracy. In addition, the platform was able to classify human breast and lung cancer subtypes using both DNA and RNA sequencing data with 96% and 99% accuracy, respectively.

Machine learning has been gaining in popularity as a way to mine for patterns in big data, integrate data from genomic and phenotypic analyses, assist drug discovery and improve patient outcomes. At this year's AACR meeting, over 30 poster abstracts involved use of machine learning in target and biomarker discovery, as well as in predicting responses to therapy. A host of companies are now utilizing machine learning in drug discovery.

---

*"We set up this network of academic collaborators to help push the envelope a bit on the work we do."*

*Hannes Smárason, WuXi NextCODE*

---

WuXi NextCode CEO Hannes Smárason told BioCentury one of the challenges in using AI to analyze omics experiments is the need to integrate data from different types of sequencing experiments carried out at different clinical centers. In addition, it can be difficult to obtain enough patient data to adequately train the algorithms, and to define the "features" of the data the algorithms should identify and operate on.

Unlike the problem of facial recognition, where 15-18 features in a person's face -- such as the distance between eyes or height of the cheekbones -- are typically sufficient to get the right answer, Smárason said genomics datasets contain hundreds of thousands of potentially useful features.

For example, those features can include specific upregulated or down-regulated genes, clusters of co-regulated genes, relative contributions of various gene clusters to a phenotype, or the identification of driver genes within gene clusters or biological pathways.

Moreover, the complexity involved in defining a disease, or even a specific phenotype or clinical outcome, within the algorithm presents a challenge for meaningful interpretation of the data.

Smárason told BioCentury WuXi NextCode has deep expertise in both AI and biology but its major advantage over other players in the space is that it has amassed a much bigger clinical dataset. "We've got access to arguably the largest sample size in the field of genomic data and that gives us a tremendous leg up when it comes to optimizing the systems we've designed," he said.

Building that dataset has been a strategic priority for the WuXi NextCode, he said, and has been accomplished by continually expanding its network of academic and industry collaborators.

On top of the Yale University collaboration, the company has academic partners at the **Massachusetts Institute**

of Technology Koch Institute, Stanford University and Harvard Medical School.

It has also struck at least dozen partnerships with non-profits, universities and companies in four key areas: population genetics, diagnostics, precision medicine and other clinical applications (see "WuXi NextCode Galactica").

## FIGURE: WUXI NEXTCODE GALACTICA

WuXi NextCode Genomics Inc. maintains a constellation of partnerships to help refine and test drive the capabilities of its AI platform, deepCODE, and to gain access to diverse clinical datasets. The figure classifies partnerships formed since 2015 into five major areas of research. It is conducting basic and translational research with academic partners at four U.S. universities, and population genetics, diagnostics development, precision medicine and other clinical applications with a variety of non-profit organizations, companies and universities across five countries.



## NEW FACE OF FGF

In the Nature study, WuXi NextCode demonstrated how deepCODE can uncover new mechanisms important for drug discovery by pinning blood vessel growth, a process that helps fuel tumor growth, to an unexpected FGF-dependent mechanism.

"We set up this network of academic collaborators to help push the envelope a bit on the work we do," said Smárason.

While FGF was known to promote growth, differentiation and survival of cells, as well as to maintain vascular integrity and prevent endothelium-to-mesenchymal transition, it was not known to be a primary driver of vessel growth or to play an appreciable role in cellular metabolism.

Smárason said the Yale team, led by Michael Simons, "had a lot of data" showing dual knockout of **FGFR1** and **FGFR3** impairs proliferation of endothelial cells as well as growth and branching of blood vessels, but enlisted WuXi NextCode to help flesh out the mechanism. Simons is a professor of medicine and cell biology at **Yale School of Medicine**.

DeepCODE analysis of RNA-seq data from endothelial cells either stimulated with **FGF2** or lacking **FGFR1** suggested FGF signaling played an unexpectedly large role in regulating metabolic gene expression, in particular for genes involved in glycolysis, a metabolic pathway relied on heavily by cancer cells.

Of the six genetic pathways most strongly linked to FGF signaling, four were metabolic.

Simons told BioCentury that while the group had suspected some metabolic pathway involvement in vessel growth and had previously identified the metabolic enzyme **HK2** in the process, WuXi NextCode's contribution painted a much more complete picture of the scope of metabolic involvement and FGF's role in it. "It would have taken years of work using conventional methods," to come to the same conclusions, Simons said.

The Yale team then experimentally validated deepCODE's predictions and showed FGF controls glycolysis through MYC-dependent expression of HK2.

Smárason believes the collaboration is a good example of how computational and biological methods can complement each other. "We were able to apply our sophisticated analytical tools to help identify a mechanism that helped explain the issue at hand. Then the Yale group was able to take that prediction and validate it using biology. It was an interesting soup to nuts example of the power of these types of methods."

The collaboration is continuing, and Smárason said the partners are working on a follow up to the *Nature* study that contains results that are even "more interesting and profound than this initial result."

Yale previously patented use of HK2 inhibition to treat heart diseases and did not file any new IP related to the *Nature* study.

Simons said patents on HK2 inhibition in cancer had already been issued to others.

Vidac Pharma Ltd. has VDA-1102 ointment, a selective modulator of VDAC1 and HK2, in Phase I testing for skin cancer, as well as ointment and IV forms of the compound in preclinical development for a variety of other cancer indications. vTv Therapeutics Inc. and partner Calithera Biosciences Inc. also have an HK2 inhibitor in preclinical development for cancer.

---

*"The importance of a general observation like this is that it starts to point you in a direction in biology that you otherwise may not have thought possible."*

*Hannes Smárason, WuXi NextCODE*

---

## I, DEEPCODE

According to Smárason, the work on deepCODE began in earnest after the subsidiary was acquired in 2015. The company's roots go back to the Icelandic genomics company **deCode genetics Inc.**, which spun out **NextCode Health** when it was acquired by **Amgen Inc.** in 2013. NextCode was acquired by New WuXi Life Science Ltd., then **WuXi PharmaTech Inc.**, in 2015 to become its third subsidiary.

While Smárason would not disclose how many patients' worth of data the AI platform has been trained on, he said the scale of the dataset has made the platform "orders of magnitude" faster and more accurate than those of competitors.

He said deepCODE can integrate diverse types of high throughput omics data, including RNA sequencing and arrays, whole-genome and exome sequencing, and metabolomics.

The program harmonizes the data, then selects features to tell the program what to look for.

"Then you set the algorithm loose and it starts this non-linear method of trying to identify patterns in the data that provide you with an accurate disease classification if that's what you are studying, or drug response," said Smárason.

He told BioCentury that as well as classifying the tumors into the 27 TCGA cancer types, the platform used a hypothesis-independent analysis of the tumors to identify a single molecular signature that predicted survival across all TCGA cancer types. He said one surprise was that there may be "a mechanism that connects all cancers."

"The importance of a general observation like this is that it starts to point you in a direction in biology that you otherwise may not have thought possible," he said.

As another example, he cited an unpublished study on classifiers for models of estrogen receptor-positive and estrogen receptor-negative breast cancer that revealed a set of genes involved in ear development were highly important. "We have no idea why that is the case," said Smárason. "The point here being the AI, the classification and the genes that come up are sometimes truly counterintuitive."

WuXi NextCode is also using its AI tools to help an undisclosed pharma partner with a regulatory problem it encountered when one country stopped paying for one of its drugs because the therapy only worked in about 5% of the patient population for which it was approved. Using deepCODE, WuXi NextCode homed in on which patients will respond to the therapy with "100% accuracy," he said, adding the pharma partner is currently in discussions that "might lead to the country re-approving the drug."

"It's one very crisp example of how you can apply this technology to have a fundamental impact both commercially and medically," said Smárason.

WuXi NextCode curates its own ontology databases and employs an iterative process that continually uses the results of its latest AI analyses to re-assess the structure and inter-relatedness of its ontology terms, further improving the accuracy of its models.

And while machine learning has now been embedded in essentially all of its computational analysis tools, the next step is to fully automate the analytics, said Smáráson. The goal is to further improve accuracy by taking human error out of the equation, and speed up all applications of the platform from target and biomarker discovery to diagnostics development and predictions of drug responses.

#### COMPANIES AND INSTITUTIONS MENTIONED

American Association for Cancer Research, Philadelphia, Pa.  
Amgen Inc. (NASDAQ:AMGN), Thousand Oaks, Calif.  
Calithera Biosciences Inc. (NASDAQ:CALA), South San Francisco, Calif.  
Harvard Medical School, Boston, Mass.  
Massachusetts Institute of Technology, Cambridge, Mass.  
New WuXi Life Science Ltd., Shanghai, China  
Stanford University, Stanford, Calif.  
Vidac Pharma Ltd., Jerusalem, Israel  
vTv Therapeutics Inc., (NASDAQ:VTVT), High Point, N.C.  
WuXi NextCode Genomics Inc., Cambridge, Mass.  
Yale School of Medicine, New Haven, Conn.  
Yale University, New Haven, Conn.

#### TARGETS

FGF - Fibroblast growth factor  
FGF2 - Fibroblast growth factor 2  
FGFR1 (CD331) - Fibroblast growth factor (FGF) receptor 1  
FGFR3 (CD333) - Fibroblast growth factor (FGF) receptor 3  
HK2 - Hexokinase 2  
MYC (c-Myc) - v-myc myelocytomatosis viral oncogene homolog  
VDAC1 - Voltage-dependent anion channel 1

#### REFERENCES

Li, V. "Next in line." *BioCentury* (2017)  
McCallister, E. "Hopes in the machine." *BioCentury* (2017)  
Yu, P., et al. "FGF-dependent metabolic control of vascular development." *Nature* (2017)