

The New York Genome Center

Nancy J. Kelley is founding executive director of the New York Genome Center (NYGC), an independent, nonprofit research organization whose mission is “to transform genomic research and healthcare for New York and the world”. Institutional founding members include Cold Spring Harbor Laboratory, Columbia University, Cornell University/Weill Cornell Medical College, The Jackson Laboratory, Memorial Sloan-Kettering Cancer Center, Mount Sinai Medical Center, NewYork-Presbyterian Hospital, New York University/NYU School of Medicine, North Shore-LIJ Health System, The Rockefeller University, and Stony Brook University. NYGC chose Illumina as its founding technology collaborator. *Nature Biotechnology* talked to Kelley about NYGC’s plan to become one of the largest genomics and bioinformatics facilities in North America.

What convinced 11 of New York’s major biomedical institutions to work with you?

New York is a major global biomedical force with unparalleled intellectual resources to support bioscience activities. Yet in the mid-1990s, Boston surpassed New York in NIH [US National Institutes of Health] funding, a gap that continues to widen. New York has also lagged with respect to the development of large-scale, centralized resources in genomics, sequencing and bioinformatics. NIH funding requirements and budgets are also getting tougher and more specific, making it more difficult for an individual investigator or an individual institution to compete at the size and scale required to become a global leader. Today, no one institution can go it alone to compete for funding and make a true transition from basic to translational research. NYGC was developed to meet those challenges head-on, creating an innovative approach to collaboration that would benefit each of the founding institutions while also creating a resource for New York and the world. Ultimately, the members agreed to contribute to NYGC because of the transformative potential for this collaboration.

What advantages arise from pooling resources?

By bringing its founding institutions and technology companies together in a unique collaboration, NYGC is able to create economies of both scale and intellect. Economies of scale include the capabilities to process full runs faster and more cost-effectively by combining orders and purchase amounts, to achieve

quicker turnaround times with operational expertise and efficiency, to implement more precise quality controls associated with creating a standardized product and to provide access to large-scale grants that might not be available to individual institutions. The entire sequencing fleet will be replaced every few years, ensuring that the latest technologies are available to all member institutions, while minimizing cost and risk. Alongside this physical infrastructure, NYGC will be positioned to recruit and develop a cross-functional team of researchers, technicians and clinicians. Furthermore, economies of intellect arise from the distributed expertise that can be harnessed to create a centralized hub of high-throughput sequencing and bioinformatics knowledge not possible at a smaller facility. The collaborative structure also allows the institutions to participate in a rich scientific ecosystem, ranging from joint research to various networking opportunities. By uniting and promoting the scientific community, NYGC will make New York an even greater magnet to attract global talent.

What core competencies will be within the center?

Facilities will include high-throughput sequencing labs; a bioinformatics group for interpreting sequencing data; a data management infrastructure that will use local and cloud-based technology to provide seamless access to sequencing and bioinformatics data and analysis for two years; an independent translational research capability led by a scientific director, with the scale and capacity to apply for NIH center-level grants; a unit focusing on spinoffs and commercialization, which will work with private sector collaborators to create new technologies, diagnostics and therapeutics; training and conference programs to provide education in genomics technology, interpretation, and applications; a philanthropic development group; and an innovation center for testing emerging sequencing platforms, beginning with Illumina, Life Technologies and Oxford Nanopore. Our own internal research, set to commence shortly, will include numerous internal teams led by senior investigators. Scientists at each of the founding institutions (some with joint appointments at NYGC) will also participate in research and grants, working groups and networking opportunities. We intend to build a team of researchers who are seasoned users of the service platforms we use, which will in turn drive the constant improvement of our services and operations.



NYGC’s Nancy Kelley says, “By uniting and promoting the local scientific community, NYGC will make New York an even greater magnet to attract talent globally.”

What are your plans for handling and storing sequence data?

NYGC will use several data centers, including an onsite data center in our primary facility, a co-location facility in Manhattan and a disaster recovery site in Washington state. The onsite center will support a data storage and computing environment dedicated to production pipeline analysis—capturing instrument data and performing primary data crunching. These systems typically need to be physically close to the sequencing instruments to simplify handling the large volumes of primary instrument data. Having them in-house also serves to protect the facility from downtime due to network interruptions beyond our control. The data center co-location site will provide very high levels of network connectivity, stable power and cooling, and other forms of stability. This location will house the long-term data archive and high-performance computing resources. NYGC will leverage Manhattan’s high levels of fiber optic connectivity to provide dedicated links between the co-location site, the genome center, ‘cloud’ computing offerings and selected founding institutions. NYGC has committed to providing 24 months of data storage, as well as access to computational and analytical resources, bundled with our primary sequencing offerings. This will require a highly scalable computing and data storage environment. Expanded analysis data will be kept only as long as they are needed for the pipeline to run. Data for long-term storage will be compressed. Both

the storage and HPC [high-performance computing] requirements are not insignificant but are attainable through multiple proven technologies and vendors.

What types of sequencing services will you offer?

We believe the three key requisites for researchers today are first, high-quality sequencing with both a timely turnaround and a competitive price, second, interpretation of raw sequence data and third, computational infrastructure, given the large size of the data generated by every sequence. NYGC offers an integrated genomics package that combines those three areas. Every project begins with an initial consultation process, where customers collaborate with sequencing and bioinformatics personnel to tailor the services provided based on the project in question. Currently, we offer whole genome, whole exome and mRNA sequencing. Over the course of the next year, we will also start offering custom targeted applications and sequencing of small RNAs, methylation, chromatin immunoprecipitation and others, as well as sequencing non-human genomes.

What about clinical sequencing applications?

At present, all of NYGC's services are for research applications. But we are committed to advancing the promise of genomic medicine and increasing the availability of sequencing for clinical use. New York has the most diverse patient population in the world making it an ideal place to include underserved populations in clinical trials. Annually, our participating academic medical institutions have 1.2 million hospital admissions, 3.7 million emergency room visits and 17 million outpatient visits. We expect that clinical genomics applications will grow over time, especially given the ongoing initiative to open a Clinical Laboratory Improvements Amendments (CLIA)-certified laboratory within the NYGC facility. We anticipate having a CLIA-approved facility up and running by late next year, which will likely become a central core of our operations. We have designed our facility such that the size and configuration of the labs may be easily changed, and we have specifically built in the capacity to expand the CLIA lab as its operations scale up. NYGC is collaborating with institutional members NewYork-Presbyterian Hospital and North Shore-LIJ [Long Island Jewish] Health System and experts in its participating institutions to develop this effort. In

addition to the regulatory and logistical challenges that will be met to open this facility, NYGC will be working with its members to develop innovative genetics counseling, data sharing and security approaches that ensure proper patient consent and privacy protection and address other ethical issues.

How do you plan to collect and store samples for sequencing?

We are developing an integrated solution that will ultimately cover all processes spanning bench to bedside, beginning with tissue banking and storage and culminating with new diagnostics, therapeutics and therapies. To that end, NYGC is actively looking to partner with a biobanking company to provide the tissue banking capability for our internal research as well as that of our founding institutions. Through our scientific working groups and ongoing discussions with New York State healthcare and scientific policy makers, we will facilitate dialog around patient consent and other ethical issues as technology continues to advance.

What about data sharing between the center and the institutions?

The plan is to work closely with our founding institutions and other partners to identify synergies. Over the coming months, we will initiate working groups to establish plans for how the respective clinical labs will work together and create a model for sharing of clinical data. At present, this is largely uncharted territory, especially given New York's stringent regulatory environment. NYGC is committed to being an innovator and leader in this space, and one of our goals involves the dedicated pursuit of change in that realm. We hope to serve as a positive influence, through discussion and cooperation with New York State healthcare and scientific policy makers, for reducing the regulatory barriers to pursuing this important research and ultimately advancing clinical care.

What about training programs?

NYGC is committed to serving as a fulcrum for genomics training, leveraging the expertise of its founding institutions and developing its own programs and courses through a dedicated training center. We will carry out courses, lab rotations, internship programs, conferences and lecture series. For example, several genomics courses are planned for next year where we work with scientists from

founding institutions. In addition, we are currently kicking off a collaboration with Cold Spring Harbor Laboratory to co-sponsor a joint medical genomics conference in Manhattan in 2014. Additional training can be provided online as part of NYGC's growing web-based presence, a strategy designed to create true thought leadership and to facilitate and lead the global discussion around sequencing and genomics.

What do you see as commercial opportunities for the center?

NYGC will offer its integrated genomics solutions to pharmaceutical and diagnostic companies, as well as co-develop its own unique solutions and processes. These relationships will range from traditional fee-for-service arrangements to corporate memberships or more extensive research collaborations. We will also provide incubators and investment funds to spur the creation of spinouts and successful ventures. Subsidized space, equipment and managerial support (including entrepreneurial education) will be made available to facilitate technology commercialization. Our innovation center will be our first such incubator-type space.

Will you be competing with commercial sequencing services?

NYGC's business model is very different from most current commercial human sequencing services. Our emphasis is on the integrated genomics solution, encompassing not only the sequencing itself but also the data interpretation and storage, which is a full package not generally available at other genome centers. In addition, the proximity of NYGC and the collaboration-based scientific fabric it will generate right here in New York set it apart from commercial companies in this space.

Do you see NYGC being replicated in other cities?

Although New York is perhaps the most obvious example of a locale where this model would work, we believe—and are hopeful—that NYGC may spur similar efforts in other cities, both in the United States and across the globe. Any city where multiple institutions are able to share resources and talent could be a potential opportunity for a new genome center of this kind. Indeed, NYGC already has been contacted by organizations around the world seeking to develop similar offerings.



© 2012 Nature America, Inc. All rights reserved.

