Navigating the Odyssey of Big Genomics in Biopharma R&D

By George Asimenos, PhD, Chief Technology Officer, DNAnexus

INTRODUCTION

In Homer’s Odyssey, Hermes gives Odysseus an acetylcholinesterase inhibitor derived from a plant called ‘moly’ to protect him from the witch-goddess Circe’s intoxicating deliriants. While modern-day drug discovery lacks the drama of Greek mythology, we are still faced with the odyssey of how to deliver on the promise of genomics-based personalized medicines. Delivering on this promise will require overcoming challenges related to the scale and diversity of datasets, the need for integrated multimodal analytics and comprehensive analysis tools, stringent controls on data security and regulatory compliance, and the critical need for collaboration across institutional firewalls.

According to the Tufts Center for the Study of Drug Development, biopharmaceutical companies are embracing this odyssey and have nearly doubled their R&D investment in personalized medicines over the past five years, and expect to increase it by an additional 33 percent in the coming five years. Some 42 percent of all drugs in development today are “personalized medicines,” and that number jumps to more than 70 percent when we just look at oncology drugs.

At DNAnexus, we’re coupling proven expertise in biopharmaceutical enterprise deployments with agile co-development to help customers navigate the genomic medicine odyssey.

EMBRACING DIVERSITY AT SCALE

In order to accelerate medical discovery, the integration of genomic and phenotypic data needs to happen on a massive scale. The DiscovEHR collaboration between Regeneron Genetics Center and Geisinger Health System, for example, brings together next-generation sequencing (NGS) data with longitudinal electronic health records for the discovery of genetic variation important to understanding human disease and therapeutic response. Two large-scale studies from this collaboration were recently published in the journal Science and discussed the analyses of genetic and phenotypic data from more than 50,000 patients. This is one of the richest resources available for drug discovery.

Regeneron is also embarking on a joint project with GlaxoSmithKline (GSK) and the UK Biobank, where the genetic variants of 500,000 people will be analyzed. AstraZeneca announced a similar initiative, which will sequence more than 2 million genomes over the next 10 years to inform discovery and development.

Apart from drug discovery, combining large cohorts of deeply phenotyped individuals with their genomic data is useful in a wide...
range of applications, including the development of companion diagnostics, patient stratification, and clinical trial recruitment.

At DNAnexus, we believe that the realization of precision medicine will require the intelligent integration of genetic data with a range of additional data types. Successfully doing this requires the technology and experience that will enable the effortless collaboration around large, diverse volumes of data in a secure and compliant fashion – which is happening on the DNAnexus Platform today.

**KEY HIGHLIGHTS:**

- **GLOBAL CLOUD** Including operations in China, Frankfurt, and North America
- **DATA COMPLIANCE** Certified ISO 27001 compliant, HIPAA, CLIA, GCP, EU regulations, & more
- **AUDIT TRAIL** All data, tools, & workflows are tracked & controlled to ensure auditability & reproducibility
- **DATA CONTROL** Through admin-defined access and permission-based controls

**WORKING TOGETHER, ACROSS BORDERS (AND HALLWAYS)**

We are now witnessing the globalization of pharmaceutical drug discovery and development. It is no longer possible to successfully operate in a single location: data (e.g. NGS and EHR) is sourced from biobanks; researchers are distributed across many locations around the world; and clinical trials can now be conducted in locations such as China. Given the complexity and scale of these efforts, infrastructure requirements must accommodate seamless communication and integration demands inherent to cross-sector partnerships as well as the legal requirements for protection of patient privacy and regulatory oversight.

Two great examples of modern-day collaboration include the Pistoia Alliance and precisionFDA. The Pistoia Alliance, in which DNAnexus is a member, supports collaboration across pharmaceutical companies in a precompetitive space, bringing together key opinion leaders to identify causes of R&D inefficiencies and develop best practices and technology pilots to overcome obstacles. PrecisionFDA, a cloud-based community platform launched in 2015 as part of the Precision Medicine Initiative (now All of Us), supports the development of novel analysis tools and regulatory science, and allows pharams to collaborate with the regulators.

This transparent approach to the global sharing of data, tools, and methodologies reflects an awareness within the genomics community that only through collaboration will the industry overcome significant challenges in the modern-day drug discovery and development odyssey.

**CREATING A BEST-IN-CLASS WORKBENCH**

A significant amount of time in the drug discovery odyssey is spent creating data analysis workflows that can integrate data, process them and generate insights that advance development. There are hundreds of tools that will answer specific questions but none that are a perfect fit for everyone. Researchers need a flexible platform where they can build a collection of best-in-class workflows and deploy their own proprietary tools.

The DNAnexus Platform integrates seamlessly with other upstream and downstream system components, including lab information management systems (LIMS) and reporting software. Hardened for enterprise-based data management, researchers can facilitate consistent workflows and ensure uniform analytical treatment and reproducibility.

M2Gen adopted the DNAnexus Platform to support data analysis and collaboration for the Oncology Research Information Exchange Network (ORIEN) Avatar Research Program. This innovative program joins academic cancer centers and pharmaceutical companies in their efforts to study and treat cancer through the development of more precise treatments for patients.
ENSURING DATA PROTECTION AND REGULATORY COMPLIANCE, GLOBALLY

Data security is not a nice-to-have when it comes to sensitive data, it is essential. With genomic sequencing emerging as a central component to clinical development and the delivery of both diagnostics and therapies, compliance with regulations that apply to the handling of these data, and its subsequent integration into other medical data systems, are equally critical. As raw data are converted into more meaningful information, they become an asset as valuable and sensitive as any other personal information, currency, or intellectual property.

DNAnexus has always taken a proactive approach to security and compliance. We’ve worked closely in partnership with Amazon Web Services (AWS) and Microsoft Azure in North America, Europe and Asia to provide our mutual customers best-in-class security in the cloud. Platform features such as two-factor authentication, end-to-end encryption, need-based network access control, 24/7 security monitoring and updates, audit and access logging allow us to provide our customers with the industry’s most comprehensive security and privacy framework.

CONCLUSION

In 2016, Frost & Sullivan (F&S) named the DNAnexus Platform one of the most significant enabling technologies helping usher in the era of large-scale genomics projects that are expected to expand to millions of genomes.

DNAnexus contributes to the modern-day odyssey by providing an environment where data is accessible in a secure and compliant manner, ultimately accelerating drug development and other pharmaceutical science.

CASE STUDIES: CO-NAVIGATING PRECISION MEDICINE WITH DNANEXUS

DNAnexus, and its broad network of partners, provides pharmaceutical and biotech companies with a secure and compliant infrastructure and the scientific support to solve today’s genomic challenges faster and more effectively than ever before. The company is the platform supporting some of the largest genomic sequencing projects in the world.

Regeneron Genetics Center: Enabling one of the largest scale integrations of next-generation sequencing and patient EHR data. DNAnexus worked closely with Regeneron to bring them online and rapidly scale up to more than 1,000 exomes processed per week and allow the seamless and secure interactions with collaborators like Geisinger Health System.

precisionFDA: The DNAnexus Platform is powering precisionFDA, providing the underlying cloud-based compute and data management infrastructure. In addition, DNAnexus is working with the FDA to build a community around its informatics platform to help drive standards around secondary analysis, the process of mapping, alignment, and variant calling of DNA sequence data.

ORIEN Cancer Research Network: M2Gen adopted the DNAnexus Platform to support data analysis and collaboration for the Oncology Research Information Exchange Network (ORIEN) Avatar Research Program. This innovative program joins academic cancer centers and pharmaceutical companies in their efforts to study and treat cancer through the development of more precise treatments for patients.