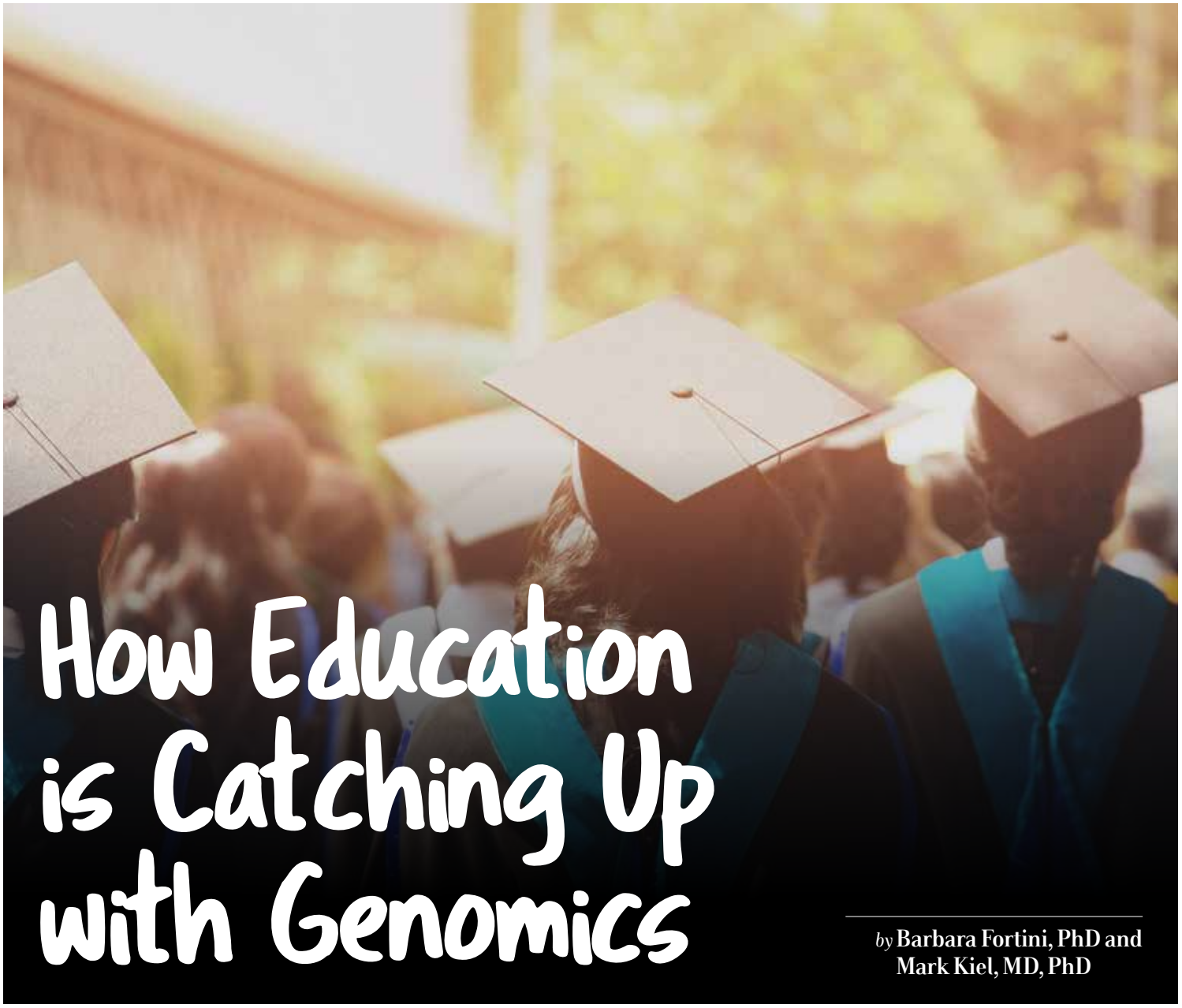

How Education is Catching Up with Genomics

by Barbara Fortini, PhD and Mark Kiel, MD, PhD

Background and Motivation

In 2017, while the world was becoming increasingly dominated by vast amounts of genomic data, genetics education programs in the US were quickly becoming at risk of being left behind. To unlock the full potential of precision medicine, educators came to realize that a new breed of geneticist was required – one that could understand and analyze this seemingly boundless information. Genomic data has been hyper-proliferating in research labs and clinics since the Human Genome Project and the subsequent precipitous drop in sequencing costs. As a result, the most critical need shifted to training a workforce fluent in analytic and computational techniques for genomic data.

> *"Equipping the next generation of genomic scientists with tools to expedite accurate data interpretation is critical to meeting the growing demands of genomic medicine"*
>
> Dr. Mark Kiel

The Promise of Precision Medicine: From Possible to Practical

The goal of the Human Genome Project (HGP) was ambitious and exciting – "to provide researchers with powerful tools to understand the genetic factors in human disease, paving the way for new strategies for their diagnosis, treatment and prevention."[1] Taking over a decade to complete and with a price tag over $3 billion, the monumental accomplishment of the HGP to sequence and map the genes within the genome made precision medicine possible in the genomics era.

With the first draft human genome complete, geneticists set their sights on a new goal – the $1,000 genome ($1KG). Scientists and physicians

KGI's Master of Science in Human Genetics and Genomic Data Analytics is a first-of-its-kind program. The first students in KGI's Master of Science in Human Genetics and Genomic Data Analytics will graduate in May 2020.

understood that making genome sequencing routine in the clinic would facilitate a powerful new diagnostic technique and enable the development of targeted therapeutics. Illumina met the goal in 2014 when they began offering whole genome sequencing for just $1,000. By 2017, DNA tests were available for over 10,000 conditions,[2] and research studies with genetic content were being published at an estimated rate of over 500,000 new studies per year. The promise of precision medicine had been made *possible* by the HGP and *plausible* by the $1KG, but what could make it *practical*?

The bottleneck had moved from *generating* data at a reasonable price to *finding, analyzing, and using* the data to help patients. Conventional analysis tools and strategies are costly, time-consuming, and subject to error-prone manual searching, annotation, and organization. Accurate diagnosis of patients is sometimes only possible after searching through the millions of scientific articles to find a single reference that makes a diagnosis possible. Not only is it time-consuming and challenging to identify and manually curate this information, there are simply not enough professionals with experience in genetics and bioinformatics to fill the current workforce demand. This gap is bound to grow with the inevitable increase in genomic data in the years to come.

> *"We were excited by the prospect of this new career of variant scientist, and recognized a great opportunity was at hand that our students could quickly fill while addressing an urgent and critical need at the forefront of applied genetics and genomics"*
>
> Dr. Barbara Fortini

Investing in the Future of Personalized Medicine: One University's Role in Revamping Genetic Education

Now that thousands of patients are having gene or genome sequencing offered every week, the need to interpret these genetic test results has led to the rise of a new career path in genetics: the variant scientist. The variant scientist is responsible for interpreting the pathogenicity of individual genetic variants according to the standards established by the American College of Medical Genetics & Genomics (ACMG) and the Association for Molecular Pathology (AMP). Their interpretations are used by genetic counselors and medical geneticists to diagnose and treat patients. The variant scientist will soon find a place in every hospital and lab as the front line between genetic data and medical diagnosis.

Excited by the prospect of this new career path of variant scientist, we (the authors) recognized a great opportunity was at hand for graduate students who could address an urgent and critical need at the forefront of applied genetics and genomics.

Taking Two Big Steps

Step 1: Establishing Genetics Degree Programs at the Keck Graduate Institute

Fulfilling this vision took two big steps. First, the School of Pharmacy (now School of Pharmacy and Health Sciences) at the Keck Graduate Institute in Claremont, CA committed to launch a Master of Science in Human Genetics and Genetic Counseling (MSGC) program. The need for licensed genetic counselors continues to

MASTERMIND®

GENOMIC SEARCH ENGINE

THE WORLD'S MOST COMPREHENSIVE DATABASE OF GENOMIC LITERATURE

REDUCE TURN-AROUND TIME

Quickly view clinically prioritized search results in an easy-to-use interface

INCREASE DIAGNOSTIC YIELD

Search into the complete text of the most comprehensive source of genomic literature

ACCELERATE THROUGHPUT

Eliminate the bottleneck of manual curation and stay up to date on the latest articles

MASTERMIND BASIC EDITION IS FREE. SIGN UP TODAY!

bit.ly/jpm-fall-2019

www.genomenon.com | hello@genomenon.com | +1-734-794-3075

outpace supply in the US, especially on the west coast. While designing the curriculum for MSGC, the anticipated need for clinical variant scientists in particular emerged as a future growth area, leading to the development of a companion program termed the Master of Science in Human Genetics and Genomic Data Analytics (MSGDA). Together, these programs prepare students to fill the growing demand for genomic interpretation and allow for a faster and more seamless transition to the workforce.

Since the MSGDA was a first-of-its-kind program, the curriculum needed to be created from scratch. The MSGDA curriculum planning team started with the MSGC program's strong foundation in molecular genetics, medical genetics, cancer genomics, and biochemical genetics coursework. To that was added practical programming and transcriptomics courses to give students a foundation in both Python and R programming languages. Finally, an introduction to the bioscience industry and medical diagnostic courses prepares students to work in a fast-moving industry setting.

KGI's intent was to make sure the students were thinking in a "genomics first" mindset, so the curriculum includes a full semester, 3-unit course in Human Genomics for both the MSGC and MSGDA students. "Genomics First" is changing how all of genetics is being done – rapidly moving from single gene and gene panel tests to (whole genome sequencing) WGS matched with RNA transcriptomes, epigenomics, protein biomarkers, ChIP-seq, and chromatin conformation capture data.

> *"Mastermind will give our students familiarity with reading primary medical literature and how to use professional genomic search tools. Mastermind automatically organizes scientific research by the genetic evidence in each article. It helps the geneticist get a comprehensive view of all the medical knowledge tied to a patient's DNA very quickly."*
>
> Dr. Barbara Fortini

Step 2: Evolving Genomic Education – Genomenon Educators Program

Through the development of these new programs, KGI realized that students needed more than just genetics and genomics knowledge in a specialized degree to succeed in this field – they needed to hit the ground running with practical variant interpretation skills that can take years to develop. To that end, KGI initiated the second big step, assembling a suite of software packages to use in the classroom that would help students overcome their learning curve in the course of a year.

As KGI was reviewing options to facilitate curation, Barbara fortuitously met (co-author) Mark Kiel of Genomenon at a conference. After a thorough review (see table), it was found that Mastermind Genomic Search Engine was uniquely suited to help students efficiently navigate the breadth of genetics literature available. As a result of our meeting

AMP Association for Molecular Pathology

AMP2019 ANNUAL MEETING & EXPO

ANNIVERSARY CELEBRATION

November 7-9, 2019
Baltimore Convention Center
Baltimore, MD

PRE-MEETING EVENTS
Guest Society Symposia,
Reference Materials Forum, &
Advocacy Hill Day
November 5

Corporate Workshop Day &
Molecular Pathology Outreach Course
November 6

AMP19.amp.org

and subsequent discussions, KGI became the inaugural school in the Genomenon Educators Program, which seeks to empower students to find and interpret information from complex genomic data sets. Part of the Genomenon Educators Program is free access to the professional version of Mastermind for all students in KGI's MSGC and MSGDA programs. The arrangement gives students hands-on experience of the way real-world variant scientists curate literature for variant interpretation using Mastermind.

The Future of Clinical Genomics at KGI

So, where do we go from here? KGI has plans to grow the student body size over the next several years, helping to fill the growing need for genomic scientists. There are also plans to expand the degree programs to include educational models that would cater to medical and industry professionals – those already working in the field would be able to learn relevant skills pertaining to genomics and data analysis which were lacking in their previous educational experiences. KGI and Genomenon are helping shape the future of precision medicine by training new variant scientists with the most advanced tools to address the challenges of variant interpretation, today and tomorrow. JoPM

> *"Genetic counselors are one of the top 15 fastest growing careers, with an expected job growth of nearly 30% in the next 7 years. With a median salary exceeding $80,000, genetic counselors are highly sought after and well paid. But genetic counseling is just one career path in this growing field"*

Software Packages and Computer Resources featured in the MSGDA Curriculum

Mastermind	Literature search engine optimized for genetics data
UCSC Genome Browser	Public genome annotation and data repository for alignment
ClinVar	Pathogenicity and allele frequency information
GnomAD	Genome aggregation database
OMIM	Comprehensive database of genes and phenotypes
Bioconductor	R script repository for open source genomics analysis tools
Bioconda	Python channel with a repository of genomics packages
Other tools also used include ExAC, DECIPHER, PubMed, COSMIC, dbSNP, TCGA, GTEx	

Barbara Fortini completed her PhD at the California Institute of Technology where her thesis focused on the biochemistry and genetics of genome stability. Barbara undertook her postdoctoral training at the University of Southern California - Keck School of Medicine in the department of Preventive Medicine using post-GWAS functional genomics to understand colorectal cancer risk. She then moved to the Keck Science Department at the Claremont Colleges teaching undergraduate genetics where she learned of sister-school KGI's interest in starting graduate programs in genetics. Through the process, she was appointed Assistant Professor of Genetics in KGI's School of Pharmacy and Health Sciences, and the Coordinator for the MSGDA program, working alongside Ashley Mills, the MSGC Program Director.

Mark Kiel is Founder and Chief Science Officer at Genomenon, where he oversees the company's scientific direction and product development. Prior to starting Genomenon, Mark completed his residency in Clinical Pathology in 2014 at the University of Michigan. While at Michigan, he completed a fellowship in Molecular Diagnostics and devised the informatics framework for clinical next-generation sequencing in the Molecular Diagnostics Laboratory. During his doctoral studies, he made ground-breaking contributions to the study of hematopoietic stem cells, for which he was awarded the Weintraub International Graduate Student Award and the ProQuest Distinguished Dissertation Award. While a post-doctoral researcher, he made significant contributions to the field of Hematopathology, including genomic profiling of lymphoid malignancies, for which he was awarded the Benjamin Castleman Award.

References

1. Yesterday, Today, and Tomorrow: Human Genome Project, https://report.nih.gov/nihfactsheets/viewfactsheet.aspx?csid=45
2. https://www.ncbi.nlm.nih.gov/gtr/
3. https://www.bls.gov/ooh/fastest-growing.htm

About Genomenon and Mastermind

Mastermind is the most comprehensive and up-to-date source of genomic literature. It saves time, improves diagnostic sensitivity, and is quickly becoming an invaluable tool in the geneticist's toolbox.

Mark Kiel, an MD/PhD in Molecular Genetic Pathology faced literature curation challenges when studying the genomes of blood cancers and setting up pipelines to analyze and interpret genome data. It was these challenges that led him to found Genomenon – a software company dedicated to streamlining genomic data analysis.

Mark designed the initial algorithms comprising Genomenon's now patented Genomic Language Processing (GLP) to break through this "bioinformatic bottleneck". When Genomenon developed the Mastermind Genomic Search Engine using Mark's GLP algorithms, the team's goal was to mimic the human curation process, but on a significantly faster and larger scale. Using proprietary artificial intelligence (AI) and machine learning (ML) techniques, Genomenon worked to incorporate all of the nuances that skilled variant scientists use to search, annotate and assess the literature into Mastermind. Over the years, there have been a number of literature search engines that have attempted to address the bottleneck, but none have solved the scaling challenge as successfully as Mastermind.

Genomenon's hope for Mastermind is not to circumvent the human curation process. Rather, the goal is to annotate the entire corpus of genomic literature in a way that is more easily discernible and digestible. Mastermind organizes the relevant literature according to associations between the diseases, genes, and variants cited in the articles, and by the clinical or functional significance of each variant based on the context provided in the text of each article. The search parameters available in Mastermind allow substantial personalization, but always manage to place the most relevant literature directly in front of the user. As the user base for Mastermind grows (already over 3,000 users in more than 100 countries), the benefit to the genomics community provided by the search engine will grow exponentially. We are greatly pleased that Dr. Fortini found Mastermind was the missing piece to equip her students for the work ahead of them.