CENTER FOR COMBINATORIAL GENE REGULATION

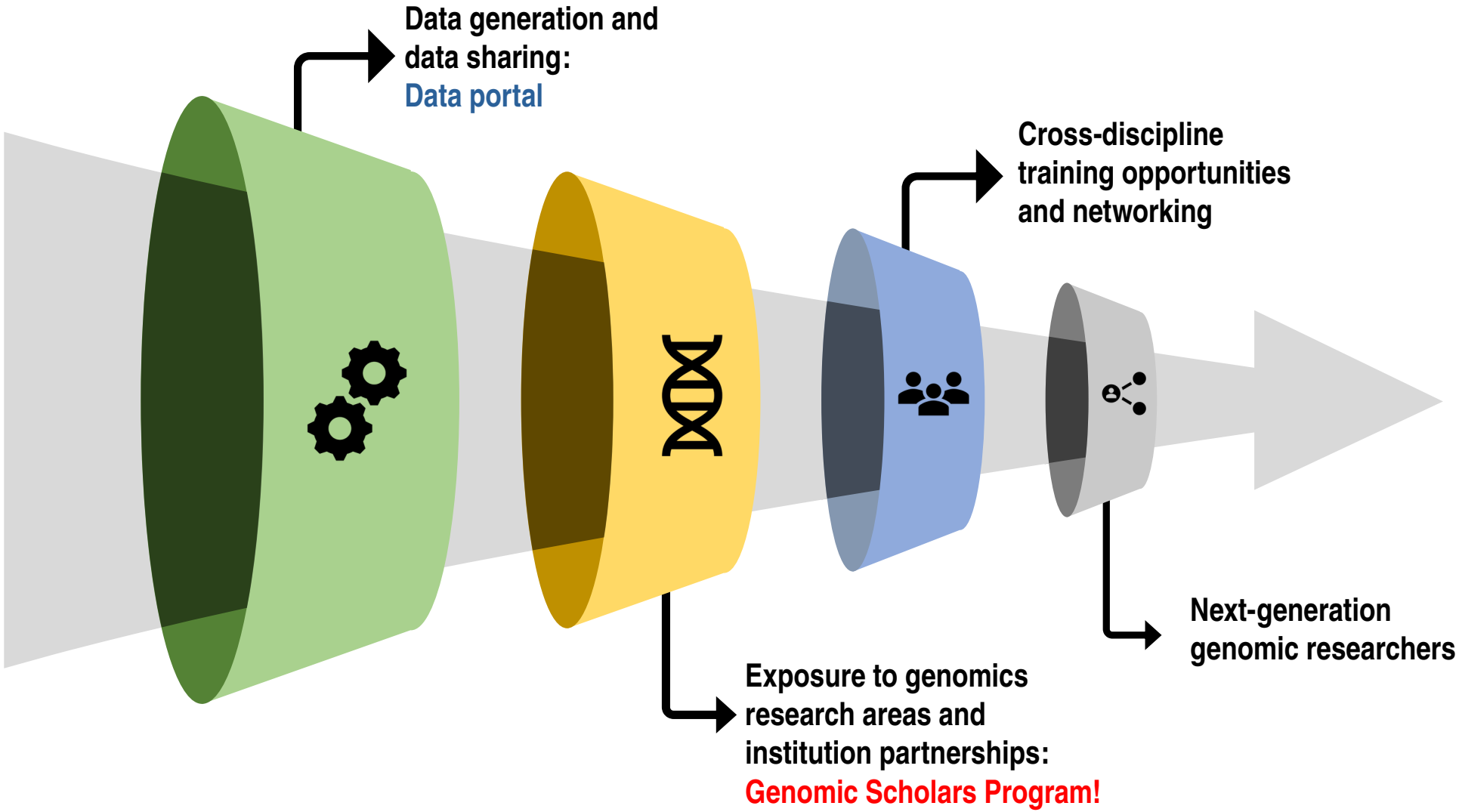
Outreach activities:
Duke Genomic Scholars Program

Alejandro Ochoa, Assistant Prof. Biostatistics and Bioinformatics, Duke University

Duke Genomic Scholars Program

https://github.com/OchoaLab/genomic-modules/

Outreach strategy



Genomic Scholars Program: Mission

Δ type	Δ score ⓘ	pre-mRNA position ⓘ
Acceptor Loss	0.84	11 bp
Donor Loss	0.00	209 bp
Acceptor Gain	0.99	2 bp
Donor Gain	0.01	-131 bp

SpliceAI screenshot

- To contribute to building a diverse genetics and genomics workforce
- Teach use of **public resources** to advance their research
- Examples:
  - Predicting the effects of genetic variants on gene regulation
  - Predicting how changes in gene regulation contribute to disease

Genomic Scholars Program lineup

Lectures

Opeyemi Olabisi, MD PhD
Nephrology

Bill Majoros, PhD
Biostatistics & Bioinformatics

Tim Reddy, PhD
Biostatistics & Bioinformatics, Biomedical Engineering, Molecular Genetics & Microbiology

Alex Ochoa, PhD
Biostatistics & Bioinformatics

Rasheed Gbadegesin, MD MBBS
Pediatrics

Allison Ashley-Koch, PhD
Nephrology

Exercises

Revathy Venukuttan
Biostatistics & Bioinformatics

Apoorva Iyengar
Genetics and Genomics

Makenzie Beaman
Genetics and Genomics

Yuncheng Duan
Biology

Shannon Clarke
Biostatistics & Bioinformatics

DID YOU KNOW...

People of African ancestry are 4 times more likely to develop kidney disease than Caucasians.

People of African ancestry have a high risk of kidney disease because of changes in the apolipoprotein L1 (APOL1) gene.

However, not all carriers of APOL1 gene changes will develop kidney disease.

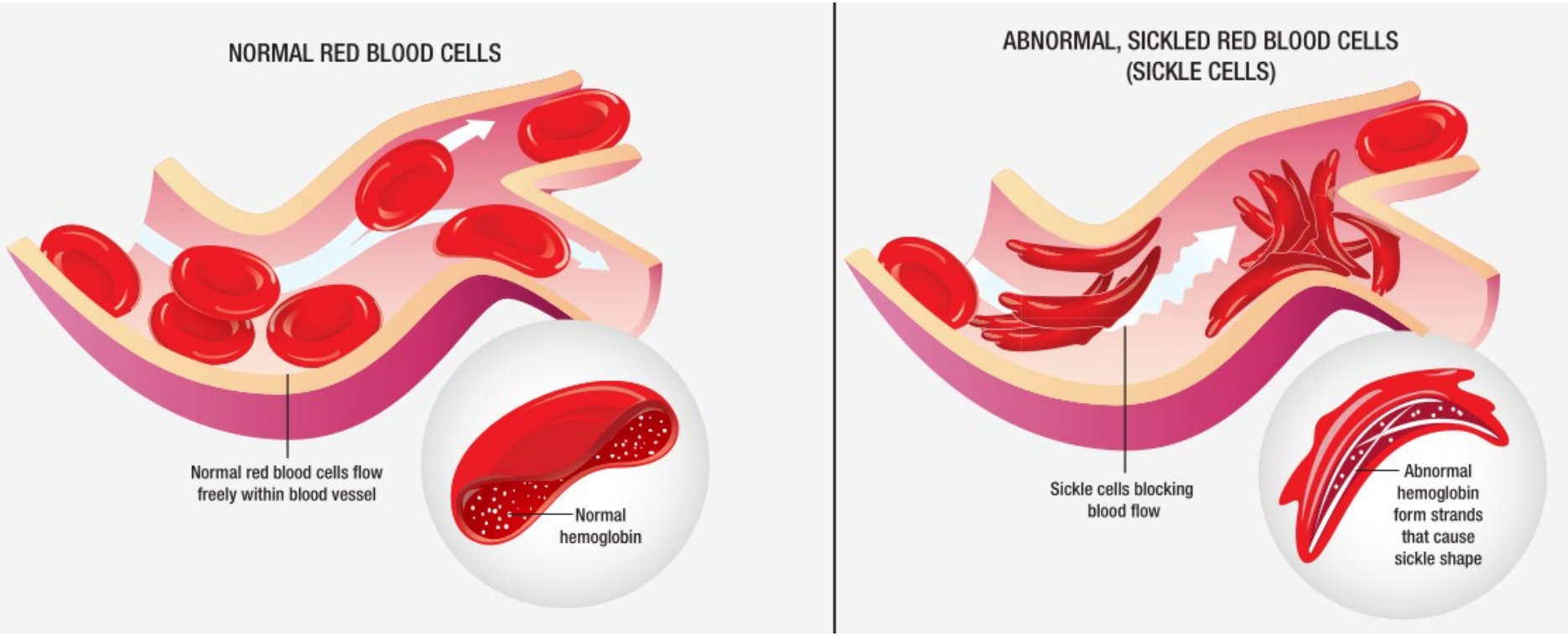
In the U.S., 13% of Blacks carry APOL1 gene changes that cause kidney disease. 70% of Blacks with diagnosis of focal segmental glomerulosclerosis (FSGS) carry these APOL1 gene changes.

Right now, there is no treatment for APOL1-associated kidney disease, and doctors don't have a way to screen for people with APOL1 gene changes who are likely to develop kidney disease.

Dr. Opeyemi Olabisi and
Dr. Rasheed Gbadegesin

https://dmpi.duke.edu/studies/apol1-study

Sickle Cell Disease



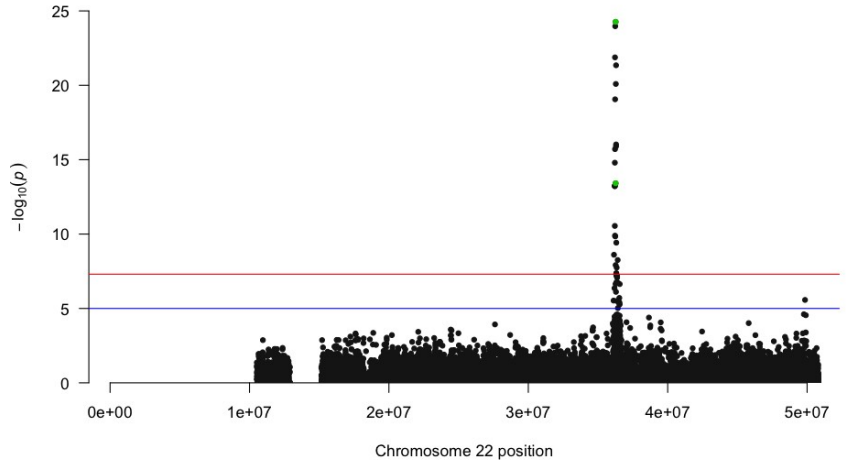
sicklecell.nhlbi.nih.gov

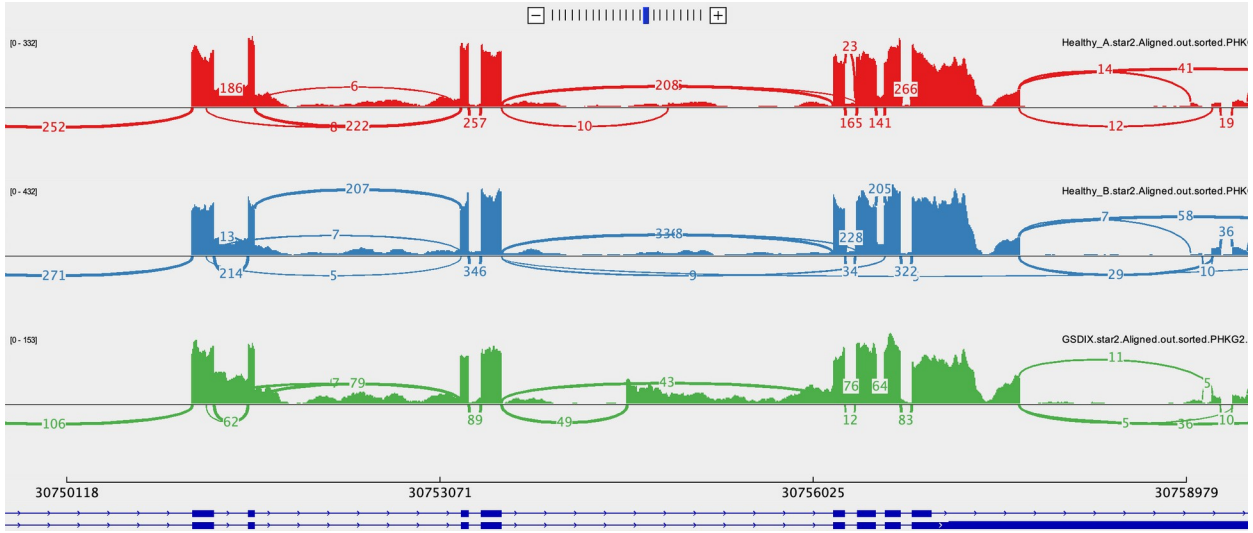
Dr. Allison Ashley-Koch

Disease allele found in gene *HBB*

BCL11A modulates disease severity: a TF with variants that turn on fetal hemoglobin!

Genomic Scholars Program: Exercises





- Built upon Data Carpentry platform
- Common thread: focus on two genetic disease/treatment loci in African ancestry (*HBB/BCL11A* and *APOL1*).
- Data, slides, other instructions publicly available on GitHub: https://github.com/OchoaLab/genomic-modules/
- Students get experience using the Integrated Genome Viewer, web tools Variant Effect Predictor and SpliceAI, and plink2 and R

Genomic Scholars Program: Next Steps

Duke UNIVERSITY

PRIME-PREP

BioCoRE

North Carolina Agricultural and Technical State University

NCCentral UNIVERSITY

NIH National Human Genome Research Institute

GENOME TDCC

The Jackson Laboratory

Check out our poster!

Duke Genomic Scholars Program: Providing Accessible Genomic Training for a Diverse Workforce

Shannon Clarke[1,2], Makenzie Beaman[3,4], Yuncheng Duan[1,2], Apoorva Iyengar[4], Revathy Venukuttan[5], Eden Harris[1], Allison Ashley-Koch[6,7], Rasheed Gbadegesin[6,7], Opeyami Olabisi[6,7], William Majoros[1,2], Tim Reddy[1,5], Alejandro Ochoa[1,2]

[1]Division of Integrative Genomics, Department of Biostatistics & Bioinformatics, [2]Center for Statistical Genetics and Genomics, [3]Physician-Scientist Training Program, [4]University Program in Genetics and Genomics, [5]Center for Advanced Genomics, [6]Division of Nephrology, Department of Medicine, and [7]Duke Molecular Physiology Institute, Duke University, Durham, NC, 27710, USA

Motivation

The genomics workforce lacks diversity and does not represent the US population. Building a diverse genomics workforce has enormous potential to improve research by fostering new ideas and approaches, and better representing the interests and motivations of the US population.

Since the sequencing of the human genome, there has been a massive expansion in the amount of freely available genetics and genomics data. Making use of those datasets (ENCODE, GTEx, gnomAD, and GWA) has the potential to dramatically lower the cost of genetics and genomics research.

Genomic Scholars Program

The **Duke Genomics Scholars Program** includes workshops offering exposure to genomic career pathways and training supporting access to dry and wet lab research opportunities.

- **Increasing awareness** of research opportunities and **providing connections** between institutions supports a pipeline of diverse representation of researchers.

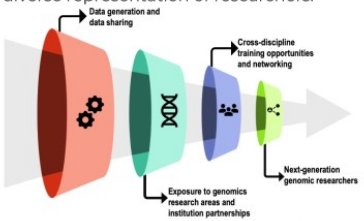


Fig. 1. Awareness, partnership, access are critical steps towards expanded diversity in the next-generation of researchers.

NIH NHGRI

Genomic Resource Workshop GitHub

Acknowledgements

This work is funded by NIH/NHGRI (RM1 HG011123).

Genomic Resource Workshop

Target Audience, Partnerships

- Interest in biology, quantitative sciences, and/or genetics and genomics
- Focus on individuals from historically marginalized communities
  - Duke **PRIME-PREP** scholars (*NHGRI funded program for post bacs*)
  - Duke **BioCoRE** (Biosciences Collaborative for Research Engagement) Program
  - Undergraduates at **local HBCUs**, including NCCU and NC A&T
- Accessible to those with limited computation/programming experience

Mode of Delivery

- Taught **in-person** and offered for a **cohort of participants** with sessions divided into focused modules that **pair lecture with hands-on exercise**
- **Common thread** across workshop with each session **building on central research question** and additional resources
- Instructor training and module format through **Data Carpentry** lesson program (https://datacarpentry.org/)

Key Components

- Highlight **career pathways** and **partnerships** with computational disciplines
- Represent **clinical and basic science** endpoints, as well as endpoints at **various levels of education**
- Inclusion of mentors and contributors who identify as individuals from historically marginalized communities
- Provide long-term access to materials and offer next steps with network for internships and rotations
- Partnership with NCCU/Duke Communication Summer Internship Program with focus to support effective recruitment efforts

Current Workshop Offering

Participants in Action

- Includes **seven PRIME-PREP Scholars** with undergraduate degrees
- Sessions led by **range of disciplines**: genomicists, bioinformaticians, physician scientists, statistical geneticists, and students from two graduate programs
- Sessions include students engaging with IGV, VEP, SpliceAI, Plink2, R, and GitHub
- Exploration of WGS, WES, and RNA-seq data files, predictors, sashimi and volcano plots
- Review hypothesis and consider research conclusions and follow-up questions

Curriculum

Type	Topic	Leader
Lecture	Biological basis of chronic kidney diseasease disparity	Opeyemi Olabisi
Exercise	Intro to IGVF and gene expression tracks	Revathy Venukuttan
Lecture	Gene structure: central dogma, splicing	Bill Majoros
Exercise	spliceAI	Apoorva Iyengar
Lecture	Consequences of variants in genes	Bill Majoros
Exercise	VEP (variant effect predictor)	Apoorva Iyengar
Lecture	Gene regulation and noncoding	Tim Reddy
Exercise	Promoter deletion In IGV	Makenzie Beaman
Lecture	Genetic association for common disease	Alex Ochoa
Exercise	Plink2 and R	Yuncheng Duan
Lecture	From Genetic Discovery to Therapy (FSGS: APOL1)	Rasheed Gbadegesin
Exercise	Pitch research questions	Alex Ochoa
Lecture	Bridging data generation, analyses, clinical interpretation	Allison Ashley-Koch
Exercise	Career Pathways	Panel

Lessons Learned

- Continuity of topic is critical in building knowledge across sessions
- Connecting lectures, hands-on exercises, research applications captures participant attention
- Identifying an individual's academic background and interest early on optimizes engagement and networking opportunities
- Highlighting the various roles of lecturers provides tangible exposure to career pathways

Shannon Clarke
Biostatistics & Bioinformatics