color

THE POPULATION GENOMICS PLATFORM Playbook

Essentials for building and scaling a population genomics program

Version 1.0
Sept. 2019

color

THE POPULATION GENOMICS PLATFORM
Playbook

© 2019 Color Genomics, Inc. *The Population Genomics Playbook* is made available under the Creative Commons Attribution-ShareALike 4.0 license.

Purpose of this Playbook

We are living through one of the most exciting times in the history of genomics. Over the past decade, we have seen a dramatic reduction in the cost of sequencing and an ever-increasing set of applications using genomic information to help prevent and treat disease.

As major health systems, payers, and national programs seek to integrate genomic information into the routine practice of medicine, a set of patterns has emerged. It is now evident that in order to offer a successful population genomics program, organizations need to address a long list of challenges. These include ensuring the analytical and clinical validity of generated data, providing tools for communication and education, managing data sharing and clinical routing, and constraining costs.

In contrast to diagnostic testing, which has been optimized to serve the 2% of the population that will need a rare disease diagnosis, population genomics programs require a different kind of infrastructure to deploy at scale. Color has worked with over 100 organizations, including health systems, large research institutions, national programs, and employers, to offer scaled genomics-based programs based on the Population Genomics Platform (PGP). We have developed the Color PGP Playbook that outlines the core architectural elements required to implement population genomics for national programs and health systems, heretofore referred to as programs. In addition, this playbook strives to reduce the problem space to a set of questions that define the core parameters that lead to a successful program.

Design Principles and Service Requirements

The Color PGP is designed to enable population genomics programs to launch quickly and efficiently by building on top of a flexible and extensible framework. It follows a set of design principles and assumes a set of service requirements, which serve as important guideposts in prioritizing capital and operational investments and in clarifying what decisions are truly material to the success of a program.

Design Principles

- **Simplicity**: The Color PGP aims to offer the simplest solution in order to satisfy the primary requirements behind population genomics programs: population health and research applications.
- **Cost effectiveness**: The Color PGP optimizes around the marginal cost per participant to provide a model that makes it financially tractable to serve entire populations.
- **Scalability**: The Color PGP is designed to scale horizontally, have minimal reliance on physical infrastructure, and use software systems to scale labor-intensive workflows.

© 2019 Color Genomics, Inc. *The Population Genomics Playbook* is made available under the Creative Commons Attribution-ShareALike 4.0 license.

- **Data architecture flexibility**: The Color PGP assumes a rapidly evolving scientific and medical landscape, which necessitates a flexible data architecture that can support updates in both the interpretation of raw data as well as the applications that are built on population datasets.
- **User-centered design:** A fundamental success factor for a population genomics program is accessibility for participants and clinicians alike. This is why the Color PGP follows a user-first design philosophy. A hallmark of successful programs is the ability to handle complex data workflows through a very simple, friendly, and easy-to-use interface.

Service Requirements

While population genomics programs have evolved from diverse early applications (such as drug discovery, clinical trials, and population health), Color has seen these use cases converge towards a unified design, with a common set of service requirements. As programs are scaling to broader populations, there is now a clear understanding that the success of any large-scale program is predicated upon the delivery of ongoing value to both the program itself and the individual participants.

- **Responsible and ethical universal return of results**: While early genomics applications integrated scientific research databases that advanced the field as a whole, benefits to individual participants were unclear. It is now sufficiently affordable and practical to deliver valuable health information to participants. For example, it is no longer excusable for a research database to house the fact that a woman carries an extreme risk for breast cancer and not afford her the opportunity to find out and pursue well-established prevention and early diagnosis options.
- **Clinical integration**: Given the important health implications of information generated by population genomics programs, these need to be implemented in such a way that findings can be effectively integrated in participant clinical care.
- **Security, privacy, and disclosure**: Genomic information can reveal significant, yet sensitive, health information. Accordingly, it is increasingly becoming a requirement for population genomics programs to comply with international and national norms governing protected health information.
- **Geographically distributed, infrastructure independent**: As programs endeavor to reach broad and diverse populations, delivery models need to accommodate a variety of access methods, ranging from point of care to fully virtualized.
- **Integrated and streamlined experience**: In a highly connected and distributed world, the success (or failure) of a population genomics program depends on its ability to efficiently capitalize on the attention, interest, and generosity of participants. Therefore, participant and clinician experiences need to be highly integrated and sufficiently straightforward to minimize confusion and loss of follow-through.

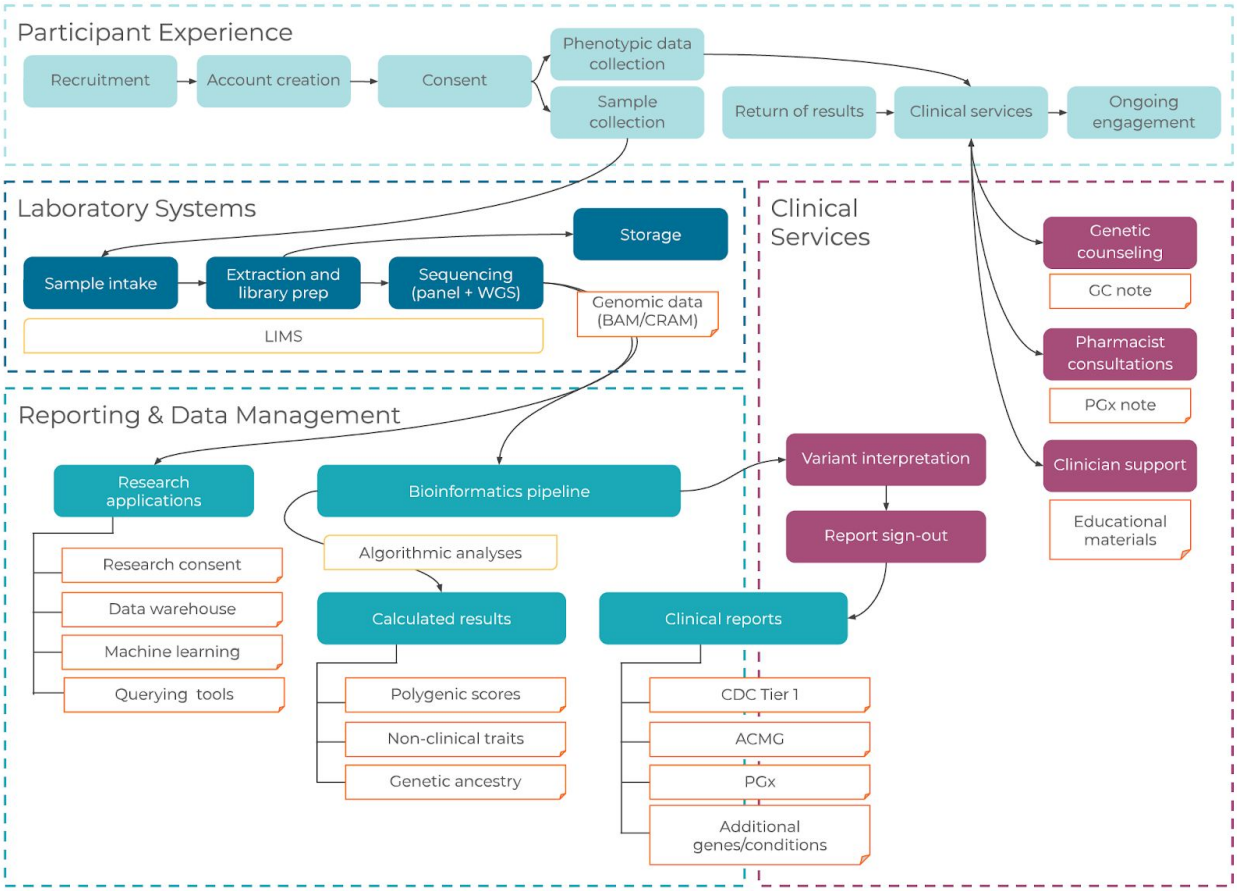
© 2019 Color Genomics, Inc. The Population Genomics Playbook is made available under the Creative Commons Attribution-ShareALike 4.0 license.

- **Durable relationships**: Population genomics programs need to be able to maintain long-term relationships with participants in order to generate longitudinal datasets as well as deliver value over time.

Architecture Overview

The Color PGP is the first deployable, fully integrated population genomics platform. It is comprised of three systems: (1) laboratory systems (sample handling, sequence generation, and secure data handling), (2) clinical services (interpretation, result sign-out, genetic counseling, electronic health records (EHR) integration, best practice alerts (BPAs), and risk cohort identification), and (3) participant experience (recruitment, consent, onboarding, participant-provided data, and ongoing engagement).

The Color PGP is designed to be deployed as a complete end-to-end solution or used modularly so that sub-systems can be satisfied by different vendors in order to augment or supplement the program's existing capabilities (such as a laboratory or clinical infrastructure).



© 2019 Color Genomics, Inc. *The Population Genomics Playbook* is made available under the Creative Commons Attribution-ShareALike 4.0 license.

Sub-System Functional Definition

Note: This playbook does not include a detailed description of each sub-system as they are well-established within the field. Rather, sub-systems are enumerated here to provide insight on the logical organization governing the Color PGP.

Sub-system	Function	Components
Laboratory Systems	• To manage biological samples and sequence these samples	• LIMS (compliance, tracking, quality control, and automation software) • Logistics (sample collection, intake, and storage) • Sample processing (extraction, library preparation, sequencing, and storage)

Clinical Services	• To manage data assets to generate clinical artifacts, provide interfaces to clinical services, and integrate with downstream data workflows (e.g. data warehouse and/or electronic health records)	• Reporting and Data Management
		  ○ Bioinformatics pipeline
		  ○ Report generation
		  ○ Algorithmic analyses
		  ○ Data workflow integration

		• Clinical Services
		  ○ Variant interpretation
		  ○ Report sign-out
		  ○ Genetic counseling
		  ○ Pharmacist consultations
		  ○ Clinician support

Participant Experience	• To manage the full participant lifecycle	• Recruitment • Account creation, authentication, and security • Consent • Sample collection • Return of results, with subsequent results updates based on new information from scientific studies • Phenotypic data collection via interactive, online tools • Ongoing engagement

© 2019 Color Genomics, Inc. *The Population Genomics Playbook* is made available under the Creative Commons Attribution-ShareALike 4.0 license.

Primary Artifacts and Data Formats

	Artifacts	Formats
Genomic data	• Targeted panel data • Whole genome sequencing (WGS) data	• FASTQ • VCF
Phenotypic data	• Phenotypic data, including social determinants of health	• Structured data (JSON) to power web user interface (UI) and EHR integrations
Clinical reports	• Center for Disease Control and Prevention (CDC) Tier 1 genomics applications: hereditary breast and ovarian cancer, Lynch syndrome, familial hypercholesterolemia • American College of Medical Genetics and Genomics (ACMG) Secondary Findings List • Pharmacogenomics (PGx) • Additional genes/conditions	• PDF • Structured data (JSON) to power web UI and EHR integrations
Calculated results	• Genome-wide polygenic scores (GPSs, using low coverage WGS, lcWGS) • Clinical risk models (e.g. Gail and Claus models for breast cancer risk) • Integrated genomic and clinical risk model scores	• Structured data (JSON) to power web UI and EHR integrations

Program Decision Matrix

Whereas the design of a population genomics program can present an overwhelming number of decisions, we have found that it is possible to reduce the problem space dramatically by initially addressing a few key choices. Once addressed, these questions yield a relatively specific set of implications, based on which a program leadership team can move into an implementation phase.

© 2019 Color Genomics, Inc. *The Population Genomics Playbook* is made available under the Creative Commons Attribution-ShareALike 4.0 license.

Parameters	Primary Options	Phase
Program goals	• Recruitment • Data asset • Market differentiation • Value-based care	**0**
Laboratory infrastructure provider	• Build • Partner	**1**
Genomic data	• 30X WGS • Targeted panel + lcWGS (for example, 1X coverage)	**1**
Clinical return of results	• CDC Tier 1 genomics applications • ACMG Secondary Findings List • PGx • Additional genes/conditions	**2**
Clinical services	• Genetic counseling (in-house or Color-provided) • Pharmacy support (in-house or Color-provided) • Variant classification (in-house or Color-provided) • Report signout (in-house or Color-provided) • Clinical services (in-house primary clinical team)	**2**
Data integration	• EHR • Research data warehouse • Clinical data warehouse	**2**
Target population	• Unselected population • High-risk population • Affected population • Ethnically diverse populations	**3**
Recruitment strategy	• Digital • Onsite • In-clinic	**3**
Data use permissions	• Clinical care • Population health management • Research and development	**3**

© 2019 Color Genomics, Inc. *The Population Genomics Playbook* is made available under the Creative Commons Attribution-ShareALike 4.0 license.

Program Goals

Different types of programs have different motivations for implementing population genomics. Primary goals may include collecting data sets for research, establishing competitive differentiation versus other clinicians, lowering the cost of care for catastrophic conditions, or improving the long-term health of a national population.

As with any complex project, it is important that population genomics programs clearly establish and gain alignment on the primary objectives. This alignment will help guide the decision-making process throughout the planning and implementation phases and lead to more successful outcomes.

One consideration that is becoming increasingly important to take into account is the social and ethical obligation to return actional results to participants. No matter what the original motivations are for a program to invest in a population genomics implementation, a consensus is emerging that it is the duty and obligation of the program to return actionable findings back to participants.

Laboratory Systems

Laboratory Infrastructure Provider (Build or Partner?)

A common challenge faced by programs in the transition from research to scaled population genomics is whether it is preferable to build a sequencing capability in-house. However, the logistics of a small-scale research or diagnostic laboratory are entirely different from those of a scaled, population genomics service.

Deciding to build a population genomics sequencing capability is a substantial investment of resources and attention. Just as software companies have increasingly found Amazon Web Services to be a better solution than purchasing their own servers, outsourcing laboratory systems may enable programs to launch and scale efficiently. However, certain conditions warrant an institution's investment of owned and operated physical infrastructure.

Below are some of the considerations that should be considered in the build vs. partner decision-making process:

- **Program goals**: Is running and scaling a genomics sequencing capability a strategic priority relative to the utilization and incorporation of genomic information into broader care delivery or research applications?
- **Security and legal parameters**: Is the program in a context that requires the primary sample processing (for example due to local regulatory requirements)?

© 2019 Color Genomics, Inc. *The Population Genomics Playbook* is made available under the Creative Commons Attribution-ShareALike 4.0 license.

- **Budgetary constraints**: Is the program budget such that it can support an upfront investment that is amortized over scale or rather one where an incremental operational cost of sequencing is more appropriate?
- **Human capital**: Is there availability of personnel with adequate skill sets required to support a clinical laboratory system, including clinical lab services, automation engineering, software systems engineering, etc.?
- **Timing**: How important is it for the program to be up and running in a short time frame relative to the 6-12 months required to get a clinical laboratory up and running?

As with most complex decisions, the build vs. partner decision is one of tradeoffs relative to priorities. Today, several organizations, such as Illumina, the Broad Institute of MIT and Harvard, and Color, provide access to scaled, cost-effective, and robust laboratory infrastructure, which may enable programs to defer or entirely forego the risk introduced by building laboratory infrastructure before launching a population genomics program. For programs that make the decision to build their own lab infrastructure, Color is happy to share its own lab design in order to help expedite the development process and reduce technical implementation risks.

Genomic Data (What to sequence?)

The Color PGP generates sequence data with three primary goals: (1) clinical-grade sequencing for clinically actionable genes, (2) population-agnostic sequencing for statistical genetics, and (3) raw sequencing data for desired option-value on future research or commercial applications.

The Color PGP sequence output includes two standard data sets:

1) **Targeted panel**: The Color PGP produces deep, clinical-grade next generation sequencing (NGS) multi-gene panel data, designed to yield clinically actionable results for population screening. This targeted panel includes high-sensitivity coverage of well-understood genes covering various clinical applications such as cancer risk, cardiovascular disease risk, PGx, carrier screening, neurodegenerative conditions, and others.

2) **WGS**: The Color PGP produces WGS data on all samples, ranging from 1X to 30X. lcWGS (for example, 1X coverage) can be used for applications of statistical genetics and achieves equal imputation quality as would be obtained from a standard genotyping array (also known as a SNP chip). As such, lcWGS can be utilized for genome-wide association studies (GWAS) as well as calculating GPSs and genetic ancestry.

   Note: Based on budgetary considerations and targeted applications, the Color PGP can increase WGS coverage anywhere from 1X to 30X or more.

© 2019 Color Genomics, Inc. *The Population Genomics Playbook* is made available under the Creative Commons Attribution-ShareALike 4.0 license.

Therefore, the decision of what to sequence in the Color PGP is reduced to the choice of depth of WGS data (between 1X to 30X), which effectively translates into whether a program has direct utility for high coverage WGS data and/or applications that rely on that incremental application.

What about whole exome sequencing (WES)?

A common question is whether it is preferable to perform WES, given that it generates a broader coverage than targeted panels and is less expensive than a clinical-grade (30X) WGS. In Color's experience, WES is inferior to the combination of a targeted panel and WGS in two critical ways:

1) For important clinical applications, an exome has insufficient depth or breadth to reliably call large rearrangements, such as copy number variants, insertions, and inversions (which are known to comprise 10-15% of pathogenic variants in high-penetrant genes like *BRCA1* and *LDLR*) . In addition, there are many difficult rare variants (such as the Boland inversion) that are clinically important and can only be detected with a targeted panel.

2) For statistical genetics applications (as described above), the WES data lacks non-coding regions, which are present in WGS data. Even in cases where WES designs add in SNPs to replicate array-based sequencing, this data remains inferior to a WGS dataset due to the ethnic bias that is inherent to SNP pre-selection (whether through array or NGS technology).

As the cost of sequencing continues to drop, it will only be a matter of time before all sequencing applications (both research- and clinical-grade) are powered by WGS data. Therefore, the utility of WES data has a limited shelf-life and is not a worthwhile investment.

What is interesting about GPSs?

Recent advancements in the field of GPSs have highlighted the ability of these scores to stratify a population into percentiles of likelihood to develop specifics diseases and phenotypes. For the purposes of population health management and disease prevention, these scores can help guide resource allocation and intervention recommendations. As part of the Color PGP, a lcWGS data is generated on every sample. This lcWGS data powers multiple applications including genetic ancestry, GPSs, and GWAS. In addition, lcWGS overcomes the ascertainment bias inherent to variant selection in genotyping array design. This can help facilitate the transferability of GPSs and other statistical genetics applications, as well as the development and validation of these applications in diverse populations.

© 2019 Color Genomics, Inc. *The Population Genomics Playbook* is made available under the Creative Commons Attribution-ShareALike 4.0 license.

Clinical Services

Clinical Return of Results (What to report?)

Genomics is becoming an integral component of clinical care and management, and the burden of determining which participants should receive clinical-grade testing, and for what genes and when, falls on the shoulders of clinicians. As a starting point for programs, the Color PGP generates three sets of clinical reports based on the latest recommendations from professional health agencies and societies such as the CDC, ACMG, the U.S. Food and Drug Administration (FDA), and the Clinical Pharmacogenetics Implementation Consortium (CPIC). These reports are revised over time to stay current with the most up-to-date evidence for variant classification, gene-set recommendations, and guidelines for care management.

- **CDC Tier 1 genomic applications:** The CDC named hereditary breast and ovarian cancer, Lynch syndrome, and familial hypercholesterolemia as Tier 1 genomics applications and estimates that nearly 2 million in the United States are at increased risk due to pathogenic or likely pathogenic variants. Several studies have found that population-level screening for these conditions leads to early detection and intervention, improved survival rates, and reduced cost.
- **ACMG secondary findings list:** In 2016 ACMG published a list of 59 clinically actionable genes that should be reported as incidental findings in clinical sequencing. These genes were carefully selected because they are associated with the more common genetic conditions and preventative measures are available.
- **PGx:** Recent studies have shown that approximately 90-99% of the United States population has at least one actionable variant in an established PGx gene. The Color PGP returns results for genes associated with a CPIC Level A gene/drug pair and with FDA-approved drug labeling.

Clinical Support Services (What specialists fulfill clinical tasks?)

A responsible population genomics program needs to take into account how to return clinical results to participants in a scalable way so that results are effectively integrated into clinical care. However, it is often cost-prohibitive and disruptive to existing clinical workflows to create and staff a separate population genomics program apart from the existing clinical staff. To this end, the Color PGP provides tools to enable programs to either staff clinical roles in-house or leverage a distributed workforce that can scale with volume.

Different clinical services are provided at multiple stages in the participant lifecycle and require different expertise and licensure. While direct care efforts must be physical and in-person, many other aspects can be remote and virtualized. The core aspects and core clinical roles required for population genomics programs are:

© 2019 Color Genomics, Inc. *The Population Genomics Playbook* is made available under the Creative Commons Attribution-ShareALike 4.0 license.

(1) Clinical interpretation and report generation
    (a) Variant classification scientists
    (b) Clinical laboratory directors
    (c) Genetic counselors
(2) Return of results
    (a) Genetic counselors
(3) Care pathway integration
    (a) Genetic counselors
    (b) Pharmacists
    (c) General practitioners (physicians, nurses, etc.)
    (d) Specialists (oncologists, cardiologists, neurologists, and geneticists)

In cases where the program is an integrated health network, all clinical services can be provided by the Color PGP and the program, with discovery of actionable results leading to direct routing of participants to the correct follow-up care.

Note: The broad adoption of a population genomics programs requires the unambiguous backing of the existing clinical staff. However, many clinicians are daunted at the prospect of having to become specialists in genetics. The Color PGP provides reference materials and resources to help educate and support clinicians.

In geographically dispersed efforts like national programs, employer-based programs, or research programs, it is important to consider what efforts will be made to support connecting individual participants with local care resources for follow-up and treatment. In these cases, the Color PGP can support secure results redistribution, providing educational materials about the program and the test results, to help optimize the ongoing participant experience.

Data Integration (Where does the data go?)

Question: How do I integrate with my clinical workflows?

To ensure that clinically actionable results are acted upon by both clinicians and participants, the Color PGP integrates results back into existing clinical workflows. Individuals at increased risk for hereditary conditions, such as the CDC Tier 1 genomic applications, should be appropriately routed for high-risk participant management. This includes routing to a high-risk cancer or cardiovascular disease clinic, where clinicians who are specialized in the care and prevention of these conditions are able to care for these individuals. Genomic results are integrated back into the EHR as discrete data fields. These data can then be used to populate BPAs which will ensure the proper integration of these types of results into the standard clinical flow.

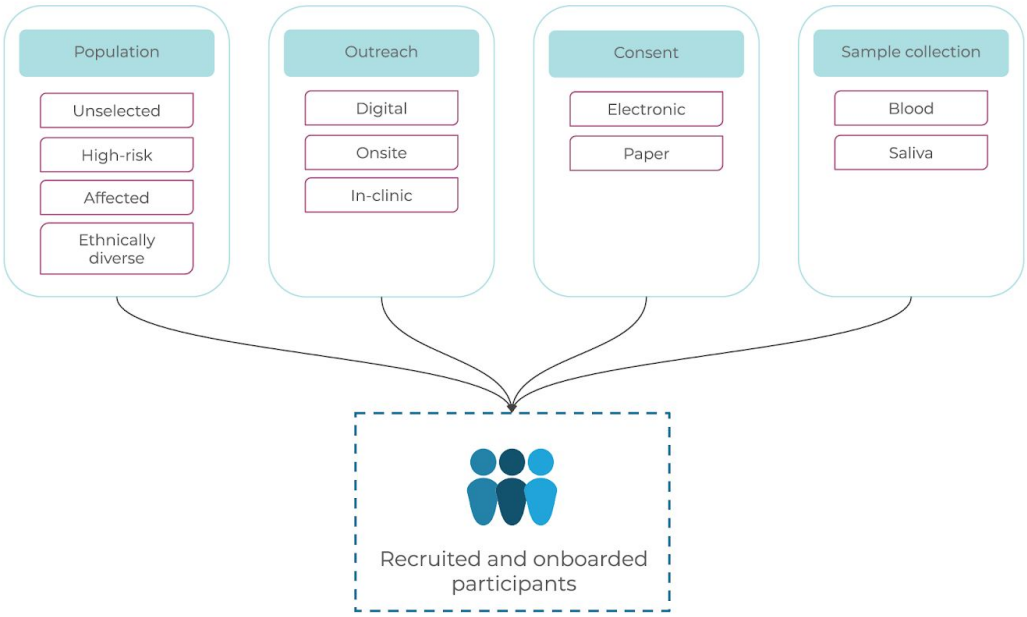
© 2019 Color Genomics, Inc. *The Population Genomics Playbook* is made available under the Creative Commons Attribution-ShareALike 4.0 license.

Question: How do I analyze the data?

A necessary side-effect of creating a large genomic data asset is simply that the data itself is large, unwieldy, and abundant. Most population genomics programs are not prepared for the sheer volume of data that is created. In order to make the data asset useful to researchers and clinicians, it is paramount that a user-friendly, powerful, and flexible data analytics dashboard is made available. As part of the Color PGP, simple data interfaces can be created (for example, data.color.com) that allow researchers to easily query, manipulate, and search the data in a fully aggregated and de-identified way to ensure participant privacy.

Participant Experience

Achieving a population genomics program at scale requires a seamless participant recruitment and onboarding experience. This includes crisp and simple educational modules that set expectations about the experience and the value of participation, clear informed consent flows, convenient sample collection, and an engaging and convenient process for gathering required information.



Target Population (Who do we recruit?)

Depending on the goals of the program, different populations are best suited for recruitment. Many genotype-phenotype databases to date have been assembled with the express purpose of discovering novel gene-disease associations or drug therapeutics. For such use cases, it is important to recruit a sufficient number of affected cases and healthy controls to power the

© 2019 Color Genomics, Inc. *The Population Genomics Playbook* is made available under the Creative Commons Attribution-ShareALike 4.0 license.

study. More recently, programs have sought to recruit and retain large unaffected (or unselected) populations for research. With the recent adoption of genomics into clinical practice, programs are now implementing genomics across their full population to inform clinical decision-making and participant care.

One aspect of existing large databases that stands to be ameliorated is the current lack of ethnic diversity. In order to ensure that genomics-based medicine is applicable to broad populations, it is imperative that new programs recruit cohorts from diverse ethnic populations.

Outreach Strategy (How do we recruit?)

It is critical to demonstrate early proof points for a large scale and ambitious program, and strong enrollment is the first key milestone. This requires clear recruitment goals with effective communication strategies that include coordinated and inviting electronic messaging, human support, and a combination of digital, onsite, and in-clinic enrollment. The Color PGP provides customizable tools for time and cost-effective recruitment using these three approaches.

Participant outreach can be driven by the population genomics program, an in-house primary clinical team, and/or the Color PGP. In digital enrollment, participants receive an email that introduces the program and prompts them to create a Color PGP account via embedded link. Once participants create an account, they learn more about the program through a 20 second video and are subsequently given the opportunity to provide consent. In onsite and in-clinic enrollment, program education and consent are supported by the program's in-house primary team. Participants can still create an account with the Color PGP by accessing a program-specific landing page.

Data Use Permissions (What do participants consent to?)

In order for a population genomics program to be successful, it is essential to effectively engage the program's many stakeholders. These stakeholders include the program itself, researchers, extended clinical care teams, and most importantly, individuals: the participants cared for by a health network or the participants in a research program.

In all cases, security, privacy, and data control are paramount. Information used for clinical purposes, even information generated as part of a research study, must be treated as sensitive. This means working with organizations that are covered entities under the HIPAA and protected under a business associate agreements (BAA) as well as minimizing the number of third-parties that have access to the data. The risk of data breach or prohibited uses increases with the number of interfaces. Implementations that rely on data passed among a collection of service providers (i.e. one entity generating data, a different entity doing clinical interpretation, a third entity providing genetic counseling, etc.) become unwieldy, in terms of implementation complexity, overhead, and data security. Simple is better.

© 2019 Color Genomics, Inc. *The Population Genomics Playbook* is made available under the Creative Commons Attribution-ShareALike 4.0 license.

A framework that protects individual privacy and permits and supports effective anonymized research and responsible identified analyses is essential. This requires careful, transparent and permissive consent and terms of service agreements, with access controls that respect the permitted and prohibited uses.

Finally, ease of use. Just as we saw with search technology and personal computing, the easier it becomes for stakeholders to interact with information, the more useful that information becomes. Simple, clean, web-based interfaces, designed around a clear understanding of the user needs, drive success. A focus on clean design, clear communication, simple analysis tools and effective visualizations magnify the personal, clinical and institutional value of a program.

Biological samples (How do we collect samples?)

The Color PGP supports collection of blood and saliva for genomic sequencing and return or results. Blood is collected by the population genomics program's in-house primary clinical team, and saliva is collected in-clinic or onsite. Alternatively, a saliva collection kit can be sent to participants' homes and be returned via mail with a prepaid mailing label.

Design Principles Deep Dives

Generating Clinical Reports

Generating a report that meets all the demands of being clinical-grade but is also accessible to the general population and easy to understand requires a robust clinical infrastructure.

Bioinformatics Pipeline

A robust clinical infrastructure begins with a bioinformatics pipeline that can be utilized for a variety of clinical NGS applications such as cancer risk, cardiovascular risk, PGx, carrier screening, neurodegenerative conditions, and others. The bioinformatics pipeline in the Color PGP has a multi-layer variant calling approach that is highly optimized for genes defined in the ACMG secondary findings list and is validated for difficult-to-call variants and regions. The pipeline also leverages variant calling and filtration methods, such as machine-learning based variant confidence models, to ensure consistent calling and processing of data.

Variant Classification

The adoption of NGS for genetic testing has brought new challenges to interpretation of sequencing variants, including an increase in the volume, type, and complexity of variants. The Color PGP provides access to PhD-level variant scientists with extensive expertise and classify variants according to the ACMG 2015 guidelines for sequence variant interpretation. To support variant interpretation in a clinical laboratory, the Color PGP also uses machine-learning approaches that weight and combine multiple levels of evidence, including computational

© 2019 Color Genomics, Inc. *The Population Genomics Playbook* is made available under the Creative Commons Attribution-ShareALike 4.0 license.

predictions of functional impact and splicing impact, location, population frequency, and aggregated individual-level information.

Reporting Infrastructure

An essential aspect of incorporating genomics into clinical care is that participants and clinicians alike understand the results and the appropriate next steps. Color PGP reports are written by genetic counselors and medical geneticists who have a deep knowledge of genomics and experience in delivering genetic test results. Within the report, results and disease associated risks are described through text and illustrations. In addition to standard static PDF reports, the Color PGP generates interactive, secure web-based reports that can be easily updated as new information from scientific studies becomes available.

Clinical laboratory directors are directly responsible for signing-out clinical reports that will be returned to participants and used by clinicians to make important medical decisions. The Color PGP Reporting Infrastructure provides laboratory directors with program workspaces that have visibility into all reportable and not reportable variants within a sample and a secure interface for individual and batch report signout. These workspaces have customizable access controls and logical isolation to ensure that each participant's data and workflow is restricted and protected.

The Engagement Lifecycle

In this section, we provide an end-to-end narrative of an effective population genomics program, to illustrate the ways a refined deployment supports the needs of the many stakeholders across the full lifecycle.

Preparing the Clinical Staff

Programs are most effective when integrated into primary care settings, but many clinicians are daunted at the prospect of having to become specialists in genetics. The Color PGP includes educational programming to ensure that the clinical staff understands the program, can communicate effectively about it to their participants, and is prepared to make effective use of the information.

The Color PGP includes department presentations to provide clinical staff with an overview of the program, refreshers on the clinical content, and common participant and provider questions. At these department presentations, support materials describing the possible clinical results and the relevant best practice guidelines are surfaced, as are the program's specific routing protocols. The key program support staff within the Color PGP are also introduced and are connected with a team of genetic counselors who can support their handling of specific cases via phone or email as they arise.

© 2019 Color Genomics, Inc. *The Population Genomics Playbook* is made available under the Creative Commons Attribution-ShareALike 4.0 license.

Return of Results

A core part of the value proposition to participants is the return of clinically useful and personally engaging information. Many programs underestimate the challenges of return of results and stumble in tackling the logistics of clinical interpretation, reporting, and participant communication at scale. A long delay in return of results can cause significant drop off in the momentum of the program.

Within two to four weeks of enrollment in the Color PGP, participants receive an email indicating that their results are ready. Participants are then routed to a website and may book a phone-based genetic counseling appointment at their convenience. Importantly, 99% of participants are able to find appointment times within 48 hours. At the time of the appointment, a genetic counselor calls the participant's phone number, carefully walks through the results, and helps schedule an in-clinic appointment with a clinician. Participants may also schedule a follow-up call with a genetic counselor, should they have additional questions. The Color PGP routes documentation of the session to the participant record in the EHR, along with the test results.

When the participant visits the clinic, the healthcare team is prepared with the results, personal and family health history information that the participant provided during onboarding, the documentation of the genetic counseling session, and information about personalized screening. If desired, the Color PGP can implement BPAs within the programs EHR to effectively utilize the PGx information that has also been generated and deposited into the participant record.

Tracking Program Progress

Visibility into the performance of population genomics programs is also essential, and the Color PGP provides dashboards to provide a high-level overview of key program metrics. Executive and business operations teams can review program effectiveness, calculate ROI, and consider opportunities for companion programs to build off the data that has been generated. They access the program dashboards to visualize recruitment rate and program follow-through, statistical summaries of the characteristics of the participants, including demographic, geographic and phenotypic information, as well as aggregate summaries of clinically actionable results and PGx impact across the full population. The team can identify practice differences between clinics and arrange a cross-pollination event so the practices of the most effective clinics can be disseminated across the network.

© 2019 Color Genomics, Inc. *The Population Genomics Playbook* is made available under the Creative Commons Attribution-ShareALike 4.0 license.

Ongoing Engagement

There are two fundamental reasons why ongoing participant engagement is essential:

1) Our understanding of the clinical and personal utility of genomics continues to expand, grow, and change. In order to ensure that participants are getting the most up-to-date and precise recommendations for care and prevention, they must continue to engage with the program to receive updates.

2) The first versions of population genomics programs were created as static, stand-alone data assets. As the field has continued to evolve, it has become not only necessary, but imperative, that the data asset continues to grow and be updated over time. Data left at-rest becomes stale. In order to ensure that the data is kept up-to-date, it is necessary to achieve high ongoing participant engagement with phenotypic surveys and other data collection mechanisms.

To achieve ongoing engagement with participants, the Color PGP offers a set of web interfaces and interactive modules, based in engagement genomics (such as non-clinical traits and genetic ancestry) and interactive surveys to ensure the ongoing collection of data as well as to capture the attention of participants.

Conclusions

As population genomics matures, several consistent elements and best practices have emerged across successfully launched programs. Color has outlined these processes in this playbook with the hope that it will serve as a helpful guide to programs building and scaling population genomics.

The takeaways from this playbook are as follows:

- Population genomics programs involve a great deal of complexity, but this complexity can be reduced to a few simple choices.
- Each program should not build every part of the genomics process from scratch. This “reinventing the wheel” approach leads to unnecessary costs, complexity, delays, and potential program instability.
- Programs should focus effort and risk on aspects that are most differentiating and most closely aligned with the fundamental program goals.
- Ethics and expectations have changed for genomics research, and it is now feasible and cost-effective to return results to research participants. Responsible genomics research should include return of results to all participants as a standard part of the research process due to the large potential health impact for every individual.

© 2019 Color Genomics, Inc. *The Population Genomics Playbook* is made available under the Creative Commons Attribution-ShareALike 4.0 license.

For more information, or to provide feedback or suggestions to improve this playbook, please email Color at populations@color.com.

Note: The Color PGP and this Playbook are distributed without warranty of any kind, either expressed or implied. The responsibility for the interpretation and use of the material lies entirely with the recipient or user. In no event shall Color Genomics, Inc. be liable for any damages or liabilities arising from its use, re-distribution, or reproduction.

::

© 2019 Color Genomics, Inc.
Except as otherwise noted, the Color PGP and this document, *The Population Genomics Platform Playbook*, is made available under the Creative Commons Attribution-ShareALike 4.0 license. To view a copy of this license, visit: https://creativecommons.org/licenses/by-sa/4.0/

"Color" and the Color logo are registered trademarks of Color Genomics, Inc. Color reserves all rights with respect to such trademarks, and such trademarks are not subject to the Creative Commons license referenced herein. Modified versions of the Color PGP and/or this playbook must be clearly identified as modified and are not endorsed by or affiliated with Color.

© 2019 Color Genomics, Inc. *The Population Genomics Playbook* is made available under the Creative Commons Attribution-ShareALike 4.0 license.