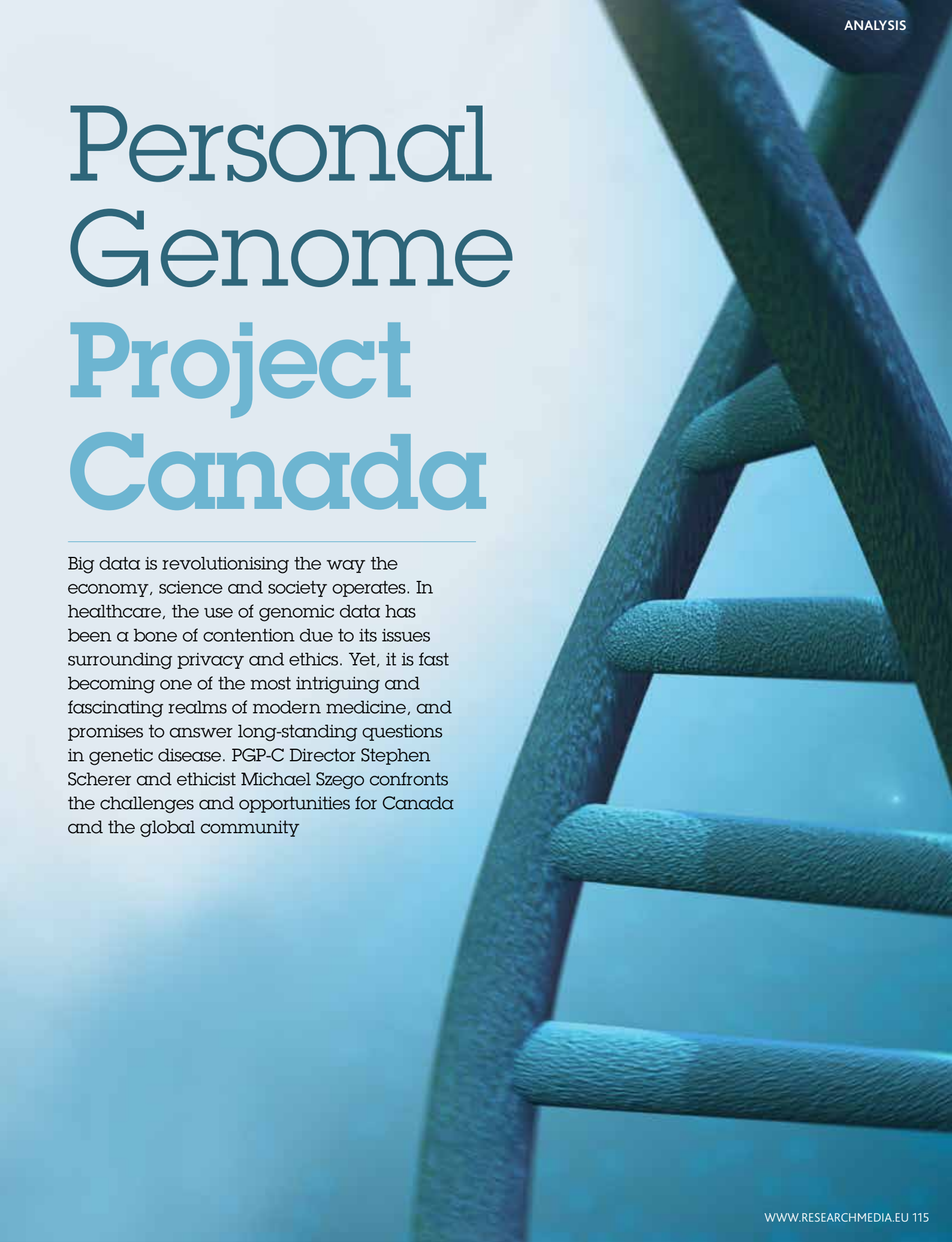
Personal Genome Project Canada

Big data is revolutionising the way the economy, science and society operates. In healthcare, the use of genomic data has been a bone of contention due to its issues surrounding privacy and ethics. Yet, it is fast becoming one of the most intriguing and fascinating realms of modern medicine, and promises to answer long-standing questions in genetic disease. PGP-C Director Stephen Scherer and ethicist Michael Szego confronts the challenges and opportunities for Canada and the global community



Could you provide some insight into your background and interest in bioethics? Has your work within the Centre for Clinical Ethics helped prepare you for your current position?

MS: I do not have a typical background for an ethicist. My doctoral training was in molecular genetics, where I studied the molecular mechanisms underlying inherited photoreceptor degenerations. I was always interested in reading ethics articles in scientific journals, but it was after attending a lecture given by a hospital-based clinical ethicist that prompted me to explore clinical ethics as a career. Upon finishing my PhD, I promptly went back to school and completed a Master's in Health Science in Bioethics and then a clinical ethics fellowship. Given my scientfic training, I am particularly interested in evidenced-based practice and ensuring the ethics debate keeps pace with scientific advances.

I think it was my background that best prepared me for my current role since I had a good understanding of both the scientific utility and the ethical issues associated with the project. My training also gave me a good working knowledge of research ethics guidelines in Canada, which was helpful when designing the Canadian protocol.

What do you hope to achieve?

SS: As is often the case in science, quantum advances in technology (in this case DNA sequencing) outpace the ability to understand what the impact of the advances might be. In essence we see the PGP-C as much of a social experiment as it is a scientific experiment, but we need each component to be grounded in the other. PGP-C is our effort to try to learn what all the questions are we should be asking before they sneak up and nip us in the heal. If we get most of the answers right we believe genome sequencing will become a mainstay in medical management, so long as the information is used in the right way

How did you come to be involved with the Personal Genome Project-Canada (PGP-C) and what does your role entail?

MS: During my graduate training in bioethics, we were required to complete a 'capstone project', which entailed practical application of all the knowledge in bioethics we had attained. I heard that Dr Stephen Scherer at The Hospital for Sick Children/University of Toronto wanted to start a Canadian PGP, inspired by the PGP-Harvard founded by George Church. I contacted Steve to see if I could be of some assistance and my capstone project became an ethical analysis of the PGP and how such a project could be adapted to the Canadian setting. I made some specific recommendations about how the project should be structured in Canada and when the course was finished, Steve suggested that I continue to work on the project and translate my analysis/recommendations into reality. After consultations with experts in privacy legislation and research ethics I adapted the PGP-Harvard protocol and consent forms to our context, which I submitted and later received research ethics board approval (the Canadian equivalent of a institutional review board). Currently, I co-facilitate all the consent conversations with research participants in collaboration with the PGP-C genetic counselor.

The PGP-Harvard was established in 2005 and has been described as a pilot group within a global network. The PGP-C was formally established in 2012 and was the first project to join the global network after the PGP-Harvard. Early work on the PGP-C, however, began back in 2006 and many of the ideas arose from Steve and his teams work on the first personal genome sequence, namely that of Craig Venter. The PGP-C, like its US counterpart, was founded on the principles that genomic information is inherently identifiable and that making genomic information publically available will act as a catalyst to improve tools for obtaining and interpreting genome information.

Because genomes are both identifiable and predictive, many research studies withhold data. How does PGP differ in its approach?

MS & SS: The overarching goal is to create a fully consented, unrestricted repository of genomic data that is linked to human trait data. Privacy restrictions can make it difficult for individuals outside of the scientific establishment to obtain access to genomic data, particularly genomic data linked to health information. However, our database can be accessed by anyone, from a scientist working at an academic institution to a computer programmer trying to find new ways to display/analyse genomic information. A massive number of genomes are required to fulfill the potential of personalised medicine. Our data will give researchers another genomic resource to utilse.

Since open access in the context of personal health information and personal genome sequences is relatively new, we employ a multi-step consent process such that participants fully understand and appreciate, to the greatest extent possible, the implications

Insight into the process

Participation will provide research participants with a unique educational experience. All research participants get a research report, indicating the variants we identified in their genome. They will also have access to a genetic counsellor to put this information into context.

The process of enroling is quite involved. Prospective research participants must first sign a mini-consent form which briefly explains what is involved. If the research subject signs the mini consent then the research subject's age, name, email address and their interest in making self-reported health and genomic information publically available is obtained. Provided the research subject is over the age of 18 and has the intention to make trait and genomic information public, then they are asked to complete the full consent form and successfully complete a 27 question entrance exam. In the consent form the subject is asked to discuss their participation with their immediate family members since a participation may involve learning information that may be of importance to other family members. If the research subject has a monozygotic twin, they too must consent prior to particiaption.

Upon obtaining a perfect score on the exam and signing the consent, the research subject is also asked to fill out a baseline trait data form in order to obtain some basic medical and trait information about the research subject. If the research subject agrees some blood is drawn and they join the waiting list for sequencing.

The website is still under development so it has been Szego's job to mail interested research participants a registration packakge with all the necessary forms and arrage a face-to-face meeting. In the near future they will move to a more online format where all the information will be available online; however, they still plan to meet each research participant to discuss consent, answer any questions and obtain the blood draw.

Another important feature of the process is that a fee of CAD $4,000 is required for enrolment. There are also some cost-sharing arrangements. PGP-C recognise this limits the number of individuals who can participate, but as they currently have limited funding this is the only way we can finance it. Scherer often says this is his and Szego's 'hobby' project.

of participation. For example, before making any decisions about publishing genomic data on the website, research participants are given their DNA sequence and a research report. After viewing all their confidential data research participants can decide to publish their sequence or they can withdraw from the study. If the latter is chosen, all data associated with the individual including published trait data and unpublished sequence data will be deleted. If the publish data option is chosen, the PGP will make the sequence data available on its public website and link it to the relevant Baseline Trait Data. While the concept of open consent may be surprising to some, it is our view that the very notion of confidentiality when dealing with whole genome information is problematic. As you point out, an individual's genome is their ultimate identifier.

What have been your major successes to date, both from an individual perspective and the Project as a whole?

MS & SS: On a personal and project level, it has been very satisfying to see the PGP-C evolve from an idea into reality. We have had a steady stream of research participants fully enroll in the project, and have had over 600 other individuals email us expressing an interest in joining. Many of the lessons we are learning from the PGP-C are behind ideas and processes for other genome projects we are conducting, including disease genome sequencing.

How are the challenges you encounter evolving?

MS & SS: One year after the launch of the PGP-C we are now facing new challenges. We had approximately 20 research participants fully enrolled and processed by February and are trying to figure out ways to scale up enrolment while keeping the consent process intact. We also attempt to come up with strategies to ensure our research subject population represents the cultural diversity present in Canada. Finally, one of the main differentiating features of the PGP-C is that we offer genetic counseling. As we start returning research results we are unsure how much time and resources will be required to maintain this important aspect of the project.

With whom do you collaborate?

MS & SS: We are engaging in collaborations with scientists that have approached us wanting to examine health outcomes of our research participants and the PGP-Harvard has been very willing to share their resources in order to help PGP-C get off the ground. Other PGP sites are being established in the UK, Asia and South America, with the collective goal of creating as many publically available genomes as possible.

In 2013, the names of people whose anonymous genome profiles were published by the 1000 Genomes Project were uncovered. This demonstrated the ability to link apparently protected genomic data to identity, fuelling the debate surrounding privacy. Can the risks associated with participation be avoided?

MS & SS: We are very clear with our research participants that there can be no expectation of privacy once genomic information is uploaded to the public website. In a sense, we mitigate privacy risks by not promising privacy and making sure research participants are aware and comfortable with this fact. In reality, what I hear from research participants is not a concern about a loss of privacy, but that the loss of privacy might lead to genetic discrimination by insurance companies or employers. The launch of the PGP-C led to a discussion about the lack of genetic nondiscrimination legislation in Canada. If facilitating such discussions leads to legislation, then all Canadians will be better protected, including PGP-C participants.

What are your hopes for the future of PGP, and more broadly speaking, for personalised medicine?

MS & SS: I hope that as we grow the number of publically available genomes, that our dataset becomes a useful resource for developing personalised medicine. I envisage a day when whole genome sequencing becomes a routine part of healthcare. Patients will have better access to preventative healthcare, and that genetic information will be utilised in both the diagnosis and treatment of disease. I hope that in a decade or so that PGP-C is thought of as the project that provided many of the answers to the complex questions that lie ahead.

www.personalgenomes.ca

PROFILE

MICHAEL SZEGO

In addition to his valued contribution to PGP-C, Szego is Research Ethicist at The Centre for Applied Genomics, The Hospital for Sick Children, Clinical Ethicist, The Centre for Clinical Ethics (a joint venture of St Michael's Hospital, St Joseph's Health Centre and Providence Healthcare) and Assistant Professor in the Department of Family and Community Medicine, University of Toronto.

STEPHEN SCHERER

Known for contributions to discovering the phenomena of global copy number variation (CNVs) of DNA and genes as the most abundant type of genetic variation in the human genome, Scherer leads one of Canada's busiest laboratories. He is simultaneously Director of The Centre for Applied Genomics, The Hospital for Sick Children and McLaughlin Centre for Molecular Medicine, University of Toronto, as well as Senior Scientist at The Hospital for Sick Children and Professor of Medicine at the University of Toronto.