A privacy risk in your DNA

EurekAlert!

New policies are needed to safeguard participants' identity in genetic studies, cautions Tel Aviv University researcher

The growing ease of DNA sequencing has led to enormous advancements in the scientific field. Through extensive networked databases, researchers can access genetic information to gain valuable knowledge about causative and preventative factors for disease, and identify new targets for future treatments. But the wider availability of such information also has a significant downside — the risk of revealing personal information.

Researchers from Tel Aviv University and the Whitehead Institute of Biomedical Research in Cambridge, MA, have developed an algorithm that can identify last names using information gathered from the Y chromosome, which passes from father to son. This finding shows how genetic data can be used to compromise an individual's privacy, says Prof. Eran Halperin of TAU's Blavatnik School of Computer Science and Department of Molecular Microbiology and Biotechnology, who worked on the project with PhD student David Golan of TAU's Department of Statistics and Operations Research and Dr. Yaniv Ehrlich of the Whitehead Institute.

Using chromosomal data drawn from genetic databases, the researchers were able to identify the surname of one in every eight people from a sample of 911 American men. Sometimes other private information could even be discovered, including their geographic locations or the identities of their relatives. Published in the journal *Science*, this result should serve as a call to action for safeguards, the researchers note.

Weighing science and privacy

"Having such data is critical to scientific research, so we must look for ways to minimize the risk, including better techniques for encrypting genetic data, education for study participants and researchers, and new legislation to protect such information and prevent its misuse," Prof. Helperin says.

Although information about a person's entire genome is often available, this project sought to determine how much can be discovered about an individual using only a small amount of chromosomal data. The researchers chose to focus on the Y chromosome, which is passed down through the male lineage, because of its connection to surnames, which are commonly passed down through the male lineage as well.

Data relating to the Y chromosome of the sample of 911 American men — whose genetic information was collected through a private company — was used to search for their corresponding surnames in public databases. The researchers' algorithm

was able to identify the family name of 12 percent of the participants.

Because they were only looking for near-precise matches, this is a very conservative return, notes Prof. Halperin. A broader search would reveal a short-list of possibilities that could reveal even more identities. And with some additional details that are commonly included in study databases, such as age group or geographic location, there is a much higher chance of tracing a person's identity, explains Golan. Those with rarer surnames were also easier to identify accurately than those with more common names.

Mitigating risk

While Prof. Halperin believes there are some positive applications of these findings, such as searching for lost relatives or identifying bodies in mass disasters, there are also serious security issues to consider. Even if the genomic data is originally anonymous, it can still be used to invade an individual's privacy — and that of their family as well. Insurance companies could use this genetic information to determine if you are at higher risk for a particular illness and ultimately deny coverage, suggests Golan.

Steps must be taken to ensure that identities are secure while allowing scientists to access valuable genomic information, the researchers say. As credit cards and other forms of ID are encrypted to extract required information while safeguarding personal details, researchers must find a way to publish genetic data in a way that it maintains individuals' privacy but still has scientific value. Those who publish their genomic information, or participate in such studies, should be made aware of the implications. And new legislation concerning the maintenance of private and public databases, as well as anti-genetic-discrimination laws, should be drafted, conclude Prof. Halperin and Golan.

Original release:
http://www.eurekalert.org/pub_releases/2013-02/afot-apr020713.php [1]

Source URL (retrieved on *11/22/2014 - 4:04am*):
http://www.ecnmag.com/news/2013/02/privacy-risk-your-dna

Links:
[1] http://www.eurekalert.org/pub_releases/2013-02/afot-apr020713.php