

Risk of genome data being tied to photos on social media found to be very small

18 November 2021, by Bob Yirka

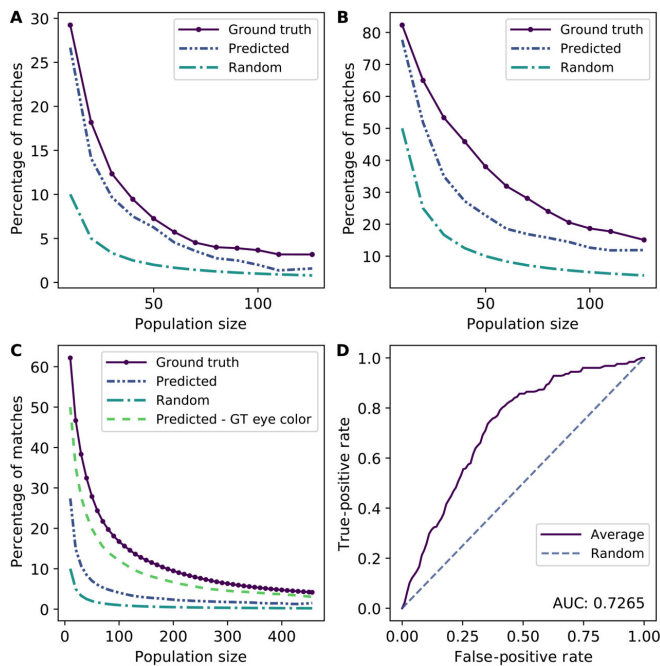


Fig. 1. Effectiveness of matching individuals' photos to their DNA sequences in OpenSNP. (A) Success rate for top 1 matching for the Real dataset. (B) Success rate for top 5 matching for the Real dataset. (C) Success rate for top 1 matching in the Synthetic-Ideal dataset. (D) ROC curve for 126 individuals. (A) to (C) present matching success results as a function of the population size (the number of individual genomes to match a face image to) for a fixed k . Credit: DOI: 10.1126/sciadv.abg3296

A trio of researchers from Washington University in St. Louis and Vanderbilt University has found a low risk for people being tied to photographs on social media sites after having their genomes analyzed by certain institutions. In their paper published in the journal *Science Advances*, Rajagopal Venkatesaramani, Bradley Malin and Yevgeniy Vorobeychik describe the ways they tested the possibility of tying publicly available genomic data to publicly available photographs.

In the recent past, some in the [media](#) have claimed that people who submit DNA samples to sites such as 23andMe are at risk of having their genomic data tied to pictures of them on [social media sites](#). The thinking has been that there is information in DNA that makes people look the way they do. And if that information could be extracted, it could be used to compare what they should look like with publicly available photographs and thereby track them. In this new effort, the researchers sought to find out if such claims are likely.

The work involved using deep learning algorithms to discover what people might look like based on their individual genetic traits. They then created a dataset with the details of 126 people's genomes and corresponding photos of them and used the same deep learning [algorithm](#) to try to tie them together.

In looking at their results, the researchers found that their system was unable to make matches in most cases—the larger the dataset, the fewer matches. But they also found that if the photographs were of high quality, the number of matches increased. They also found that the most limiting factor was the ability of algorithms to pick up eye color in publicly available photographs because of their low quality. They suggest that it is very unlikely that a person could be visually identified via their genomic data, at least given current technology. They also found that adding small perturbations into the photographs dramatically reduced the number of matches their system could find, suggesting that it would be a trivial matter to prevent tying genetic data with photographs.

More information: Rajagopal Venkatesaramani et al, Re-identification of individuals in genomic datasets using public face images, *Science Advances* (2021). [DOI: 10.1126/sciadv.abg3296](https://doi.org/10.1126/sciadv.abg3296)

© 2021 Science X Network

APA citation: Risk of genome data being tied to photos on social media found to be very small (2021, November 18) retrieved 1 December 2021 from <https://techxplore.com/news/2021-11-genome-tied-photos-social-media.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.