



Police patrol Kashgar in China's Xinjiang region, where there are reports of systematic human-rights abuses against the Uyghur population.

JOHANNES EISELE/AFP/GETTY

# CRIME-SOLVING DNA DATABASE FACES ETHICAL SCRUTINY

Geneticists say a global Y-chromosome database holds profiles from men who are unlikely to have given free informed consent. **By Quirin Schiermeier**

In May 1999, a disturbing crime shocked the inhabitants of Kollum, a small village in the Netherlands. A local 16-year-old girl was found raped and murdered in a field nearby, and some people said that Iraqi or Afghan residents at an asylum seekers' centre in the village could be to blame. Tensions rose: a fight broke out at

a planning meeting about the centre. With the case unsolved, the public prosecutor turned to a newly launched research database containing Y-chromosome profiles from men across the world. When forensic scientists compared DNA from semen collected at the crime scene with profiles stored in this Y-chromosome Haplotype Reference Database (YHRD) and

elsewhere, they found that the murderer was very probably of northwestern European descent, showing that the villagers' assumptions were unfounded. The discovery helped to calm social tensions – although the case was not solved for many years until, with the aid of more DNA work, a local farmer was found guilty.

The YHRD, which was first released online in 2000, is now widely used across the world to help solve sex crimes and settle paternity cases. Holding more than 300,000 anonymous Y-chromosome profiles, it shows how particular genetic markers are fingerprints of male lineages in more than 1,300 distinct global populations. It can point to the likely geographic origin of mystery males, as in the Kollum case, but is now more often relied on to calculate the weight of evidence against a male suspect whose Y-chromosome DNA profile matches traces found at a crime scene. Although the YHRD is a research database, scientists both from academia and crime laboratories have uploaded data to it, and it has become a key tool for prosecutors and defence lawyers.

"The YHRD is absolutely essential for suspects anywhere in the world to get a fair chance in court," says Walther Parson, a forensic geneticist at Innsbruck Medical University in Austria, and the vice-president of the International Society for Forensic Genetics (ISFG).

But some European geneticists say that the

YHRD has an ethics problem. Thousands of the profiles it holds were obtained from men who are unlikely to have given free, informed consent, they say. These include data from minority ethnic populations such as the Uyghurs in China and the Roma in eastern Europe (see ‘Populations in a forensic database’).

The criticisms raise questions about consent checks at the YHRD and other databases. They are also the latest strand of a wider campaign to draw attention to a ramping-up of DNA profiling across the world in the absence of stringent ethical oversight. Some researchers are calling for geneticists to dissociate themselves from this by ensuring that such DNA studies are not given credence by being published in academic journals or databases.

Yves Moreau, a computational biologist at the Catholic University of Leuven in Belgium, has been leading calls for journals and the YHRD to investigate potentially unethical studies. He focuses on China, where the authorities have drawn international condemnation for mass detentions and human-rights abuses in the northwestern Xinjiang province. Journalists, human-rights groups and academics have collected testimonies from the Uyghurs, a predominantly Muslim population there, of abuses that include the forced collection of DNA. (The Chinese government says it is engaged in a re-education campaign in Xinjiang to quell a terrorist movement.)

“China is pursuing, and trying to export, an authoritarian political model based on surveillance technology, including genetic surveillance of minorities,” says Moreau, who also notes a controversial Chinese effort to build a domestic forensic database by collecting DNA from millions of men across the country. “Researchers, database curators and science publishers should not be complicit in this model – absolutely in no way.”

In part because of Moreau’s work, one set of DNA profiles has been removed from the YHRD, but many more have been questioned. *Nature* has learnt that the ISFG is now setting up an oversight board to examine cases in which consent is unclear. “Some might think that ethics is merely a lot of annoying bureaucracy,” says Moreau. “But that’s a false perception. Researchers must understand that they won’t get recognition for unethical research.”

## Ethics concerns

Police forces in many nations collect DNA from suspected or convicted criminals, but usually keep this information privately for a limited time (defined by national law) and do not seek consent that would allow it to be shared publicly. By contrast, public, international DNA research databases that map human genetic diversity contain information acquired from across populations, and require that researchers who upload data have ethics-committee approval and informed

consent from participants for their studies. These databases allow scientists to calculate the frequency with which particular inherited genetic clusters exist in populations, findings that are often useful for medical research.



**IT IS EXTREMELY DOUBTFUL THAT SUCH STUDIES WERE ALWAYS DONE WITH PEOPLE’S INFORMED CONSENT.”**

The YHRD is the largest database to focus on male lineages, and is unusual for its close connection to forensics and crime-solving. It is curated by two forensic geneticists, Sascha Willuweit and Lutz Roewer, who are both at Berlin’s largest research hospital, the Charité. In common with other databases, it asks for, but doesn’t verify, consent or ethical approval. Researchers who wish to upload data to it are encouraged to first publish an academic paper about their study – which puts the onus on a scientific journal to check for ethical compliance. In about 10% of cases, says Roewer, scientists have uploaded data without publishing a paper about their work.

In March 2019, Moreau began looking into studies of minority populations in China. He spotted a 2017 review<sup>1</sup> of almost 38,000 Y-chromosome profiles of men in China, held in the YHRD. It was co-authored by Willuweit and Roewer; other authors included researchers from Chinese public-security and police forces. “I realized that the YHRD was a problem,” says Moreau. The review stated that the profiles had been collected with informed consent – but Moreau argues that it’s hard to see how Uyghurs and other persecuted minority groups could have freely given it.

Moreau went on to find dozens of articles in leading international forensics journals, co-authored by members of the Chinese police, that described DNA profiling of Tibetans, Uyghurs and other minority groups. Most of these papers were not related to the YHRD. He contacted Springer Nature, which publishes some of the journals, and it began to investigate. (*Nature’s* news team is editorially independent of its publisher.) In December 2019, Moreau published his concerns in an opinion article<sup>2</sup>, without mentioning the YHRD. Then, in January 2020, he asked the Charité to investigate. It declined, saying that it hosted but did not own or operate the database, a stance it has reiterated to *Nature*. Moreau began discussing his concerns with

Roewer and Willuweit.

Roewer says he agrees with Moreau that researchers shouldn’t be using DNA taken without informed consent – but says that it is not the YHRD’s job to check or to initiate investigations. “We are not an ethical control body,” he says. (He also says that, because the 2017 study was a review of already-collated data, it didn’t need ethics approval.)

Last year, Roewer removed profiles that had been uploaded by the authors of a paper describing DNA studies of Uyghur, Kazakh and Hui minority groups in China. He did so after Springer Nature investigated and retracted that paper; the authors said that the study had been undertaken without the approval of an institutional ethics committee. Roewer says he’ll remove related data if there are other retractions. Springer Nature has so far added editor’s notes to 28 papers – including the 2017 review – stating that concerns have been raised about their ethics-approval and informed-consent procedures. It has also retracted a second paper. “We do expect to take further editorial action in at least some of these cases,” a spokesperson says. Moreau says he has identified at least nine other potentially problematic studies – involving co-authors from the Chinese police – for which data appear on the YHRD, but hasn’t yet alerted journal editors or publishers to those concerns.

The YHRD contains eight data sets of Uyghur profiles that were uploaded directly. Roewer says that in these cases, too, it is not his responsibility to initiate investigations. “Anyone concerned about a particular data set would do better to contact the authors directly,” he says. (*Nature* e-mailed 15 researchers who’d uploaded Uyghur data; one, Yiping Hou, a forensic geneticist at Sichuan University in Chengdu, replied that the data were obtained with “valid informed consent”.)

In November 2020, the YHRD was criticized from a different angle when three researchers in Germany – science historian Veronika Lipphardt, sociologist Mihai Surdu, both at the University of Freiburg, and geneticist Gudrun Rappold at the University of Heidelberg – published preprints of their work on genetic studies about the Roma<sup>3,4</sup>. The researchers noted that the YHRD holds profiles from population studies on eastern European Roma and, in some cases, that the uploaders’ studies didn’t make clear whether they obtained informed consent, or when the profiles were collected – and that they sometimes thanked police forces for collecting the DNA. “It is extremely doubtful that such studies were always done with people’s fully informed consent,” Lipphardt says, adding that there is a long history of discrimination against the Roma.

Lipphardt tried to follow up one such concern<sup>4</sup>: in 2017, German police uploaded to the YHRD 74 profiles of male Romanians and

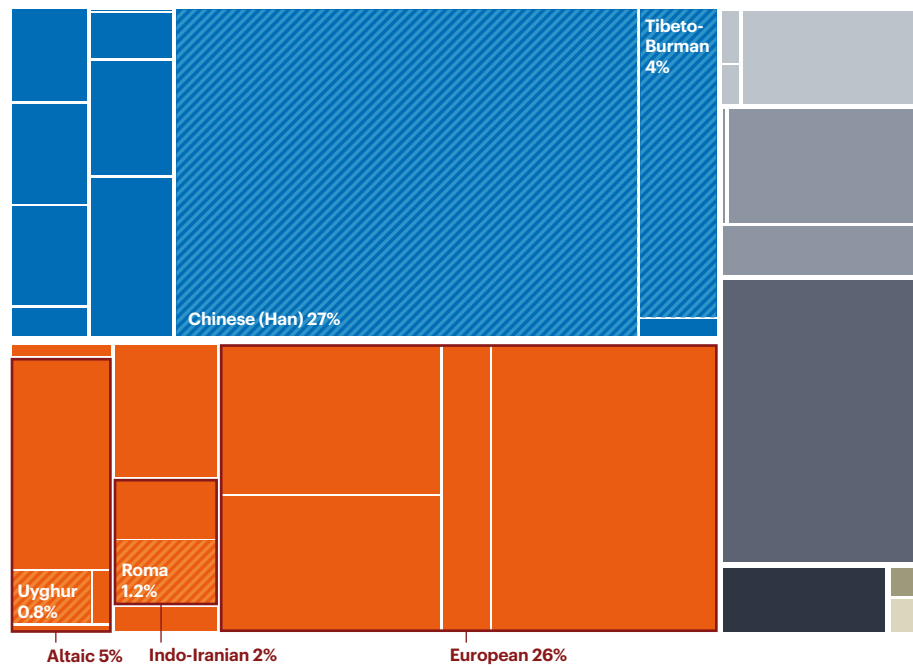


## POPULATIONS IN A FORENSIC DATABASE

The public Y-Chromosome Haplotype Reference Database (YHRD) holds information on signature markers in Y-chromosome profiles from men around the world. Some ethicists query whether Uyghur, Roma and other populations gave free, informed consent when providing data.

### Share of YHRD profiles:

East Asian 41% Eurasian 36% Afro-Asiatic 4% African 6%  
Admixed 10% Native American 2% Eskimo Aleut 0.2% Australian Aboriginal 0.2%



\*Specific populations highlighted, using YHRD-defined meta-population names. Full breakdown at <https://yhrd.org/pages/resources/stats#metapopulations>. Data as of 9 May 2021.

Afghans living in Germany. Lipphardt asked the relevant criminal investigation authority – in the state of Baden-Württemberg – whether the data were described in a paper, but was told that authorities considered the YHRD upload sufficient publication, and so no information was available on ethics procedures or consent. “Criminal investigators obviously have less stringent ethical standards than academic scientists when collecting and working with genetic information,” she says. Roewer says the criminal-investigation authority did talk to him about the issue, but it decided not to retract the data sets.

The field of forensic genetics was slow to catch up with ethical standards of biomedical genetics, Lipphardt says. It was only in 2010 that leading forensics journals introduced requirements for authors to mention informed consent or approval by ethics committees. Much of the data in forensic databases, including the YHRD, was collected before 2010, so even when data are connected to research papers, it is hard to find ethics statements.

The YHRD isn’t the only international database under scrutiny. In their preprint<sup>4</sup>, Lipphardt and Surdu note that a mitochondrial-DNA database called EMPOP also holds Roma data from journal studies that don’t explicitly state they have informed consent or ethical approval. (EMPOP is not a public database; its users must register.) Parson, who

curates EMPOP, says that all profiles undergo “rigorous quality control” including ethical evaluation before they are uploaded, and adds that the Roma data were published in scientific journals with ethical requirements to which EMPOP adheres.

Peter Schneider, a forensic geneticist at the University of Cologne in Germany, says that because the YHRD and other databases hold information only on particular genetic markers, and not full DNA sequences, individual donors can’t be uniquely identified. He thinks that in such cases, keeping data accessible is more beneficial to society than harmful to an individual.

Roewer says that when there are concerns over consent procedures for uploaded DNA profiles that were never described in journals, an independent, objective board of experts should investigate them. The ISFG is going to set one up, he and Parson say: the society is currently deciding the panel’s remit.

### Public outcry

The YHRD ethical discussion was bubbling behind the scenes, but spilled out online after Lipphardt and her colleagues published their studies, which the newspaper *Süddeutsche Zeitung* reported on. Then, in January 2021, the Charité decided to close the department for forensic genetics which maintains the YHRD – not for ethical reasons, but for economic

ones. A court ruling had required German police authorities to allow competition to provide DNA analysis for criminal investigations, which meant that the department could lose much of a guaranteed income stream that it had relied on.

The decision, made public in February, caused a huge outcry among geneticists and public prosecutors fearing loss of forensic expertise, who also pointed out how useful the YHRD was. In March, the Wie-DNA initiative, a group of social scientists, geneticists and anthropologists who analyse how DNA is used, and which includes Lipphardt, issued a statement saying it hoped the forensic institute could be saved, but also noting the ethical concerns around the YHRD’s holdings.

In April, the Charité yielded to mounting political pressure and reversed its decision – but in May, the *Berliner Zeitung* newspaper reported that researchers in the Charité’s forensic genetics institute would all be offered jobs direct with police authorities. (A Charité spokesperson told *Nature* that nothing had been decided, but that Roewer and Willuweit would preserve the YHRD in any eventuality.)

But the wider ethical debate continues. This January, Moreau and two other researchers writing on behalf of the European Society of Human Genetics published a commentary<sup>5</sup> calling for academic institutions worldwide to stop collaborations with groups over which there are ethical concerns. “These problems extend beyond China,” the commentary stated, giving the example of the YHRD. “We would like to see an end to collaborations between academic and clinical institutions worldwide and institutions in countries carrying out widespread, unethical DNA collections and/or analysis.” The authors added that any study involving police or judicial authorities in authoritarian regimes should be considered “potentially ethically tainted”.

Parson says that, although he agrees that genetic profiles obtained without informed consent should be removed from databases, the YHRD must not be curtailed more than necessary. Removing lineages of minority groups might distort analyses of the likelihood of a particular Y-chromosome profile coming from such a population. “Judges anywhere in the world rely on robust forensic data,” he says. “Excluding data from minority groups could bias statistical evaluations in forensic reports – to their disadvantage.”

Quirin Schiermeier is a senior reporter for *Nature* in Munich, Germany.

1. Nothnagel, M. et al. *Hum. Genet.* **136**, 485–497 (2017).
2. Moreau, Y. *Nature* **576**, 36–38 (2019).
3. Lipphardt, V., Rappold, G. & Surdu, M. Preprint at ResearchGate <https://doi.org/10.13140/RG.2.2.13286.04165> (2020).
4. Lipphardt, V. & Surdu, M. Preprint at ResearchGate <https://doi.org/10.13140/RG.2.2.16641.48484> (2020).
5. Forzano, F. et al. *Eur. J. Hum. Genet.* **29**, 894–896 (2021).