

# Forensic Genealogy

Serial killers and Short Tandem Repeats



SOCIETY NEWS  
--

NETWORKING NIGHT

FEATURED ARTICLE:  
FORENSIC GENEALOGY

OPPORTUNITY:  
Victor Chang  
Cardiac Research Institute

QUOTE  
VISUALISATION OF THE WEEK  
CONTACT

# Society News



# BINFSOC

## Networking Night.

BINFSOC's 2nd annual industry networking night is this Tuesday 14th June at the MCIC (6PM - 8PM)

This is where you will be able to meet leading industry and academic representatives, hang out with your BINF buddies outside of BINF2010 and make those valuable connections to kickstart your career!

REGISTER HERE:

<https://forms.gle/HrvW3uSX9FJUWXQF9>

Event:

<https://fb.me/e/1QKU3L7sb>

# Opportunity: Winter Scholarship.



**Victor Chang**  
Cardiac Research Institute



The Victor Chang Cardiac Research Institute Winter Scholarship Program provides undergraduate students with the chance to experience the life of a researcher at one of Australia's leading biomedical research facilities.

This exciting opportunity will allow you to broaden your knowledge of biomedical research, hone your communication skills and participate in laboratory activities. You will work alongside some of the most experienced and passionate researchers in the country and get an insight into what studying for Honours or other higher degree might be like.

A limited number of Winter Scholarships are available and will be awarded based on merit. Successful candidates will undertake 6 to 8 weeks of research between commencing in July - August 2022.

**more information:** <https://unswbinfo.com/victor-chang-winter-scholarship>

“

We all carry inside us, people  
who came before us.

— Liam Callanan

# **forensic genealogy**

HOW BIOINFORMATICS HAS HELPED SOLVE SOME OF  
THE BIGGEST UNSOLVED CRIMES -- AND THE DARK  
SIDE OF ITS USE

*Writers Tom Parish, Cam McMenamie*

# serial killers and short tandem repeats

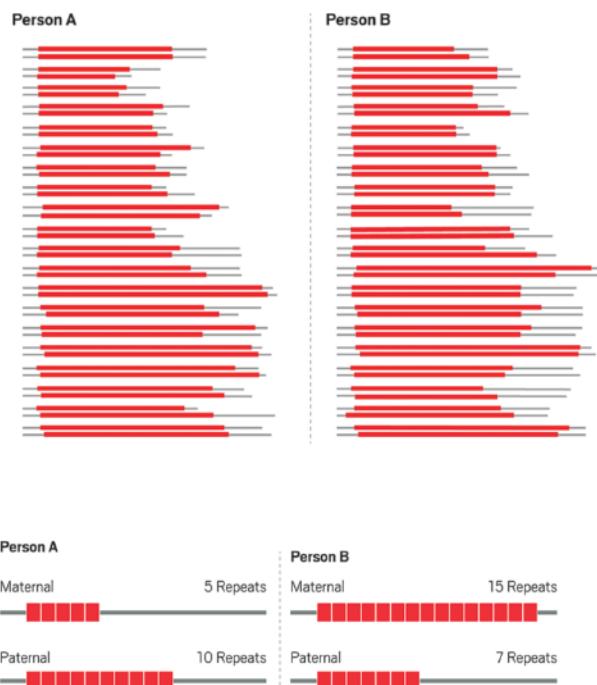
## History of DNA in forensics

DNA has been used extensively in forensics to link a crime to a suspect since the 1980s. The original technique was pioneered by British geneticist Sir Alec Jeffreys, and has been used to solve paternity disputes and assist police detectives. The principle remains the same today, where a comparison of DNA profiles is made to link a sample to an individual's 'genetic fingerprint'; but the techniques involved have dramatically improved since. In the early days, it could take 6 to 8 weeks to perform the DNA analysis. Now, it takes less than 2 hours to generate a profile after the sample is collected. The development of this technology was occurring in parallel to the use of PCR to amplify DNA (increasing the sensitivity of the test), and the commencement of the Human Genome Project, both of which led to massive shifts in our understanding of DNA itself, and pushed forward the use of DNA analysis in forensics.

## What is DNA Profiling?

Even in diverse human populations, the percentage of an individual's DNA that is shared between others is about 99.9%. Despite this, there are regions which are highly variable, termed polymorphisms. These regions vary significantly between individuals, and are determined by a unique combination of both parents' DNA. It is these polymorphisms that are used to create a DNA profile, which can not only uniquely identify you, but also give information about your relatives. The polymorphic regions used in DNA profiles for forensics are called Short Tandem Repeats (STRs). These are pieces of non-coding DNA (i.e. they are not directly translated to a protein

sequence), which include a number of repeats of the same short DNA sequence. For example, an STR might be the subsequence "GATA" repeated 6 times.



These STRs can be analysed at a number of locations, on various chromosomes to form a unique profile. The number of repetitions that each STR has is not in itself unique, but the combination of these STRs provides a unique collection of DNA sequences that can be used to uniquely identify an individual. Due to the nature of combinatorics, a sufficiently large number of STRs collected will dramatically decrease the chances of two individuals having the same profile.

The usefulness of the profile doesn't stop here. Since these highly variable genetic regions are inherited, they will also be very similar between parent and children, as well as between siblings. As a result, a family tree can be constructed

based on the information contained within these profiles.

## Distant relatives

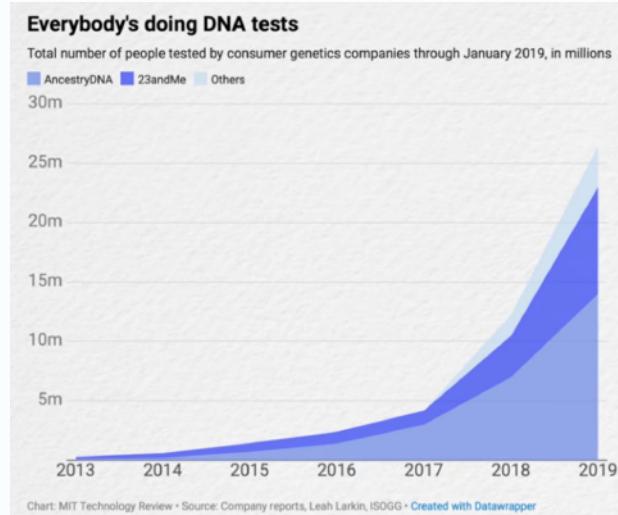
Historically, if you were trying to use a DNA profile to solve a crime, you would need to obtain a DNA sample of a direct relative, or of the suspect themselves. This was problematic, because the database of genomic information accessible by law enforcement rarely had enough information to do so. What these databases did contain were more distant relatives to the suspect, such as second or third cousins. Now, this is not helpful on its own, as there may be hundreds of people with an equivalent amount of genetic similarity to you, but if multiple of these matches are used, a family tree can be constructed to find the common ancestor and directly relate the samples to you as an individual.

This can be done programmatically, as the process is essentially building a phylogenetic tree; however, a lack of data points restricts this, and often relies on public data records and more traditional genealogy methods. This is mainly because previous generations do not have nearly as many available genetic sequences as today. The result is identical to a phylogenetic tree constructed by programs such as the beloved PHYLIP and MEGA.

So, as an investigator, you're now left with a region of the tree in which you know your suspect lies, however this still may be in a region of 1000 people. You've still got some work to do. Taking into account other 'metadata' factors such as age, height, location and ethnicity, you can narrow this down to a more reasonable collection of suspects.

As the size of DNA databases increases, the likelihood of being able to identify a 3rd cousin of the DNA sample also increases. In fact, you only need to have 2% of the population sequenced in a

database to be able to identify a 3rd cousin or closer in 99% of cases. Some services such as Ancestry.com and 23andMe do not work with law enforcement, but others such as familyTreeDNA and GEDmatch do.



In the foreseeable future, it is conceivable there are enough samples in databases accessible by law enforcement, that a program can identify an individual based on their DNA without any external input.

## Cold cases

There have been a number of high profile cases solved this way, including a number of serial murderers. One such case is that of the Golden State Killer, an infamous serial killer active from 1974 to 1986. There was DNA left at some of the crime scenes, but a lack of a match meant that the case went cold until around 2016. It wasn't until 2018, using the methods outlined above, that the suspect was narrowed down to a single family. A DNA test confirmed the suspect was Joseph DeAngelo using DNA sharing site GEDmatch.

The oldest case to have been solved using this new method was the 1956 murder of Lloyd Bogle and Patricia Kalitzke. The case had long been inactive, having occurred more than 60 years ago, but given developments in other cases such as the Golden State Killer's, detectives felt they had enough information to reopen the case. They used

genetic information to build a family tree, and narrowed it down to one suspect, and a DNA test of family members confirmed it.

These two cases show the power of this technology to identify the perpetrators of crimes that occurred more than half a century ago. This technology is being used to solve many serious crimes occurring today as well, and it's still evolving. Imagine the possibility of being able to use a DNA sample from a strand of hair or fingerprint to determine physical attributes of a suspect -- things like eye colour, ethnicity, sex, or maybe even age (base methylation levels can be an indicator of age) could be identified from the sample alone. However, this degree of access for governmental agencies has historically been an area of concern for some people.

## Privacy Concerns

As genomic sequencing technology and the bioinformatics that make it possible have exponentially advanced in the last 20 years, the laws and processes surrounding them have not caught up nearly as fast.

Additionally, while the sequencing technology that exists today is both cheap and extremely accurate (beyond 99% in NGS), the cross-referencing to ancestral records is not nearly as sophisticated, and mistakes can and do happen.

What may be more troubling than simply getting someone's hair or eye colour wrong, is the extent to which inaccurate inferences made from the original DNA sequence is used against an individual. With more encroachments made into our personal lives with social media, web-based tracking, and even in our own homes (e.g. Google home or Alexa), our genetic fingerprint is one of the few things that we have left that can be truly ours -- if we keep it that way. One can only imagine the consequences if this information becomes just another data point that is used by massive tech companies to track and

characterise us; and is a frequently explored topic in dystopian fiction such as 'Black Mirror'. Will health insurance companies charge you more for premiums if they know you have a genetic predisposition to certain diseases? What will targeted ads on social media people be like if they know who is genetically more likely to become addicted to certain things?

Perhaps technology will help protect against these threats. Some promising examples of using blockchain technology to encrypt genomic samples are already in place, which allow the data to be studied for scientific research but nullify its use in forensics and other privacy-concerning cases. This may also incentivise more people to have their genomes sequenced for science, with the peace of mind of still-intact privacy. In the meantime, we can only hope that DNA testing companies that hold our data have rigid privacy policies.

Importantly, the issue is not necessarily that there is any danger or risk right now -- but the future is uncertain. Will the companies that store our genetic data act as ethically as they do now, in the long term? Will legal loopholes be discovered that are invisible to us now that allow more and more exploitation of this data? These questions are as of yet unanswered, but careful consideration and public discourse must be undertaken to ensure the security of all our data in the future.

The key principle that is problematic here is the impermanence of the DNA record. Once you have your genetic information uploaded somewhere on a database, you can't go back. Your DNA will not change throughout your lifetime. It's what makes you up -- so once it is 'known' by an entity, in theory it will always be known. And of course, in the information age, the concept of 'deleting' data is an interesting one -- once something is digitised, it can, in theory, always be replicated: much as the original DNA molecule can.

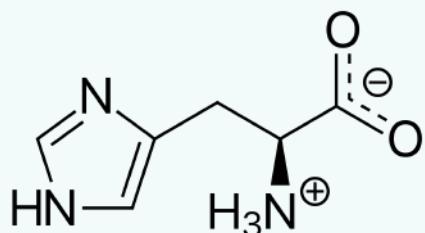
## AMINO ACID OF THE WEEK

---

[ HISTIDINE ]

CHEMICAL STRUCTURE

---



HISTIDINE

H

His  
155.1546

DNA CODONS

C A Y

POLARITY:

DEPENDS ON pH

POSITIVELY CHARGED AT PHYSIOLOGICAL pH

DISCOVERY:

ISOLATED IN 1896 BY PHYSICIAN Albrecht Kossel AND Sven Gustaf Hedin

PROTEINOGENIC - BUILDING BLOCK OF PROTEINS.

PRECURSOR TO HISTAMINE, IMPORTANT AS AN INFLAMMATORY AGENT IN THE IMMUNE SYSTEM.

CAN FORM COMPLEXES WITH MANY METALLIC IONS

IMIDAZOLE SIDECHAIN SERVES AS A LIGAND IN METALLOPROTEINS, NOTABLY ATTACHING TO Fe (IRON) IN HEMOGLOBIN (PROTEIN FOR OXYGEN TRANSPORT IN RED BLOOD CELLS)

# Contact us



IF YOU HAVE ANY COMMENTS or feedback regarding BINFsights, please write to us at  
**[binfo@unswbinfsoc.com](mailto:binfo@unswbinfsoc.com)**

We also encourage anyone to share with us anything you'd like us to take a look at, be it a bioinformatics tool that you have made or find useful; or news in the bioinformatics world that you'd like to see written about in future issues.



TO VIEW PAST AND PRESENT issues of BINFsights, check out our website at **[unswbinfsoc.com/binfo](http://unswbinfsoc.com/binfo)**

Stay tuned on our Facebook page for updates regarding events and society news.

-- The BINFSOC Team

**BINF**  
sights.