

[Request Information](#)[Permalink](#)

# Methods and Computational System for Genetic Identification and Relatedness Detection

Tech ID: 33476 / UC Case 2022-808-0

## BACKGROUND

Deoxyribonucleic acid- (DNA-) based identification in forensics is typically accomplished via genotyping allele length at a defined set of short tandem repeat (STR) loci via polymerase chain reaction (PCR). These PCR assays are robust, reliable, and inexpensive. Given the multiallelic nature of each of these loci, a small panel of STR markers can provide suitable discriminatory power for personal identification.

Massively parallel sequencing (MPS) technologies and genotype array technologies invite new approaches for DNA-based identification. Application of these technologies has provided catalogs of global human genetic variation at single-nucleotide polymorphic (SNP) sites and short insertion-deletion (INDEL) sites. For example, from the 1000 Genomes Project, there is now a catalog of nearly all human SNP and INDEL variation down to 1% worldwide frequency.

Genotype files, generated via MPS or genotype array, can be compared between individuals to find regions that are co-inherited or identical-by-descent (IBD). These comparisons are the basis of the relative finder functions in many direct-to-consumer genetic testing products. A special case of relative-finding is self-identification. This is a trivial comparison of genotype files as self-comparisons will be identical across all sites, minus the error rate of the assay.

For many forensic samples, however, the available DNA may not be suitable for PCR-based STR amplification, genotype array analysis, or MPS to the depth required for comprehensive, accurate genotype calling. In the case of PCR, one of the most common failure modes occurs when DNA is too fragmented for amplification. For these samples, it may be possible to directly observe the degree of DNA fragmentation from the decreased amplification efficiency of larger STR amplicons from a multiplex STR amplification. In the case of severely fragmented samples, where all DNA fragments are shorter than the shortest STR amplicon length, PCR simply fails with no product.

## TECHNOLOGY DESCRIPTION

To overcome these challenges, researchers at UC Santa Cruz (UCSC) invented computer-implemented methods and related software for comparing genotype data from a first sample to a limited amount of DNA sequence data from a second sample. This approach, called IBDGem, is a fast and robust computational paradigm for detecting genomic regions of IBD by comparing low-coverage shotgun sequence data against genotype calls from a known query individual. At less than 1× genome coverage, IBDGem reliably detects segments of relatedness and can make high-confidence identity detections with as little as 0.01× genome coverage. In certain embodiments, the first sample is from a known individual and the second sample is an unknown sample. The methods find use in a variety of contexts, including for genetic identity detection, e.g., for forensic and other applications. Also provided are computer-implemented methods for assessing the degree of relatedness between genotype data from a first sample and a limited amount of DNA sequence data from a second sample. Computer-readable media and systems that find use in practicing the methods of the present disclosure are also provided.

## APPLICATIONS

- ▶ forensics
- ▶ paleogenomics
- ▶ genomics

## ADVANTAGES

## CONTACT

Marc Oettinger  
[marc.oettinger@ucsc.edu](mailto:marc.oettinger@ucsc.edu)  
tel: 831-502-0253.



## INVENTORS

- ▶ Green, Richard E.
- ▶ Nguyn, Remy

## OTHER INFORMATION

### KEYWORDS

genomics, DNA identification, genome, human genome, identity-by-descent, relatedness, identical-by-descent, IBD, IBDGem, STR, short tandem repeat, forensic, forensics

### CATEGORIZED AS

- ▶ **Biotechnology**
  - ▶ Bioinformatics
  - ▶ Genomics
- ▶ **Computer**
  - ▶ Software
- ▶ **Medical**
  - ▶ Research Tools
  - ▶ Software

### RELATED CASES

2022-808-0

- Specifically designed to work with low-coverage sequencing data and where PCR cannot.
- Reliably detects segments of relatedness.
- High-confidence identity detections with as little as 0.01× genome coverage.
- Fast, robust, and scalable.

INTELLECTUAL PROPERTY INFORMATION

Country	Type	Number	Dated	Case
United States Of America	Published Application	<a href="#">20230105167</a>	04/06/2023	2022-808
Patent Cooperation Treaty	Published Application	<a href="#">WO 2023/043825</a>	03/23/2023	2022-808

RELATED MATERIALS

- [Remy Nguyen, Joshua D Kapp, Samuel Sacco, Steven P Myers, Richard E Green, A computational approach for positive genetic identification and relatedness detection from low-coverage shotgun sequencing data, Journal of Heredity, Volume 114, Issue 5, September 2023, Pages 504–512 - 06/29/2023](#)

ADDITIONAL TECHNOLOGIES BY THESE INVENTORS

- [Barcoded Solid Supports and Methods of Making and Using Same](#)

University of California, Santa Cruz  
Industry Alliances & Technology Commercialization  
Kerr 413 / IATC,  
Santa Cruz,CA 95064

Tel: 831.459.5415  
[innovation@ucsc.edu](mailto:innovation@ucsc.edu)  
[officeofresearch.ucsc.edu/](http://officeofresearch.ucsc.edu/)  
Fax: 831.459.1658

© 2024, The Regents of the University of California  
[Terms of use](#)  
[Privacy Notice](#)