Final publication is available from Mary Ann Liebert, Inc., publishers
https://doi.org/10.1089/forensic.2025.0001

Forensic Genomics

Extended Kinship Inference. Part 1: Evaluation of Short Tandem Repeats and Single Nucleotide Polymorphisms using Likelihood Ratios and Haplotype Matching

Journal:	*Forensic Genomics*
Manuscript ID	FORENSIC-2025-0001.R1
Manuscript Type:	Original Articles
Manuscript Keywords (Search Terms):	Single nucleotide polymorphism, Human identification, Targeted amplicon sequencing, Forensic investigative genetic genealogy, Kinship, Short tandem repeat
Keywords:	Autosomal DNA polymorphisms, mini- and microsatellites (or short tandem repeats, STRs), Bioinformatic methods for defining genetic relationships, single nucleotide polymorphisms (SNPs), X and Y chromosome polymorphisms
Abstract:	Medium- and long-range kinship analysis using single nucleotide polymorphism (SNP) genotyping allows law enforcement to generate investigative leads by identifying an unknown individual through their close and distant genetic relatives. Short-range kinship inference can be conducted through calculation of likelihood ratios (LRs) using population allele frequency data and the degree of similarity between two DNA profiles but has historically been limited to short tandem repeat (STR) profiles. Alternatively, identical-by-descent (IBD) segment matching algorithms can be used to detect shared DNA haplotypes between two genetically related individuals that have been inherited from a common ancestor. The ForenSeq® Kintelligence Kit enables law enforcement and forensic laboratories to utilise medium-density SNP sequencing technology for extended kinship inference by analysing 10,230 SNPs. In this study, DNA from two pedigrees was used to compare the ability of STR profiles, identity-informative SNP (iiSNP) profiles, Kintelligence profiles and Kintelligence and direct-to-consumer profiles available on public genetic genealogy databases to detect and classify genetic relationships. The DNA profiles were analysed using DBLR™ software to calculate kinship LRs or uploaded to GEDmatch PRO™ for IBD segment matching with either database searching or one-to-one comparisons. The LRs calculated for STR and iiSNP profiles were able to correctly infer first degree relationships (i.e. parent, offspring and full sibling), with the combined discrimination power able to distinguish between second degree relationships. LRs calculated for the Kintelligence profiles exceeded one million for 93% of full sibling to fifth degree relationships tested. IBD segment matching was effective for detecting first to fifth degree relatives when Kintelligence profiles were searched on the GEDmatch PRO™ database. The results of this study demonstrate that the Kintelligence Kit is a valuable tool for law enforcement and forensic investigators, offering an advanced method for medium-range kinship testing using either LRs or IBD segment matching.

Extended Kinship Inference. Part 1: Evaluation of Short Tandem Repeats and Single Nucleotide Polymorphisms using Likelihood Ratios and Haplotype Matching

Abstract:

Medium- and long-range kinship analysis using single nucleotide polymorphism (SNP) genotyping allows law enforcement to generate investigative leads by identifying an unknown individual through their close and distant genetic relatives. Short-range kinship inference can be conducted through calculation of likelihood ratios (LRs) using population allele frequency data and the degree of similarity between two DNA profiles but has historically been limited to short tandem repeat (STR) profiles. Alternatively, identical-by-descent (IBD) segment matching algorithms can be used to detect shared DNA haplotypes between two genetically related individuals that have been inherited from a common ancestor. The ForenSeq® Kintelligence Kit enables law enforcement and forensic laboratories to utilise medium-density SNP sequencing technology for extended kinship inference by analysing 10,230 SNPs. In this study, DNA from two pedigrees was used to compare the ability of STR profiles, identity-informative SNP (iiSNP) profiles, Kintelligence profiles and Kintelligence and direct-to-consumer profiles available on public genetic genealogy databases to detect and classify genetic relationships. The DNA profiles were analysed using DBLR™ software to calculate kinship LRs or uploaded to GEDmatch PRO™ for IBD segment matching with either database searching or one-to-one comparisons. The LRs calculated for STR and iiSNP profiles were able to correctly infer first degree relationships (i.e. parent, offspring and full sibling), with the combined discrimination power able to distinguish between second degree relationships. LRs calculated for the Kintelligence profiles exceeded one million for 93% of full sibling to fifth degree relationships tested. IBD segment matching was effective for detecting first to fifth degree relatives when Kintelligence profiles were searched on the GEDmatch PRO™ database. The results of this study demonstrate that the Kintelligence Kit is a valuable tool for law enforcement and forensic investigators, offering an advanced method for medium-range kinship testing using either LRs or IBD segment matching.

Keywords:

Single nucleotide polymorphism; human identification; targeted amplicon sequencing; forensic investigative genetic genealogy; kinship; short tandem repeat.

Supplementary Material: Supplementary Information S1: Calculations for expected shared centimorgan (cM), average identical by descent (IBD) segment length and number of IBD segments.

1. Introduction

Kinship inference is a valuable identification tool to generate investigative leads from an unknown DNA sample where other forensic identification methods have been unsuccessful. This has been applied to the identification of persons of interest (PoIs) in criminal investigations and unidentified human remains

in missing person investigations. The technique involves uploading an unknown DNA profile to a relevant database and collating a list of any DNA profiles that partially match the unknown DNA profile (a kinship match) that could be genetic relatives or directly comparing the unknown DNA profile to a known family reference sample.[1,2] Depending on the genetic marker panel, first (e.g. parent/offspring) to ninth (e.g. fourth cousins) relationships can be detected and classified as possible genetic relatives.

Alleles in common between two DNA profiles are identical by state (IBS). However, if the alleles are inherited from a common ancestor, they are also identical by descent (IBD).[3] Statistical support for relatedness between two individuals can be calculated using the likelihood ratio (LR) approach. This is the ratio of two conditional probabilities for observing IBS alleles in the DNA profiles under alternative hypotheses about the genetic relationship between the individuals.[4,5]

Short tandem repeat (STR) markers, the most commonly used marker type in forensic genetics and accepted in many jurisdictions as a scientific method for identification, have facilitated short-range kinship testing since 2002.[4,6,7] However, STR profiles have only been routinely applied to first degree relationships (parents, offsprings and full siblings) as more distant genetic relatives are difficult to distinguish from unrelated individuals with the limited number of markers targeted by STR panels.[8] In contrast, single nucleotide polymorphism (SNP) panels target more markers and can therefore expand the reach of kinship inference to detect more distant relationships.[3,9]

It is possible that two alleles are IBS not because they are IBD but because of mutations occurring in different lineages independently. This can be accounted for with medium- to high-density SNP panels that target SNPs in close proximity to each other on the genome. These markers are less likely to be separated during recombination and therefore co-inherited, forming a haplotype or IBD segment.[10,11] Algorithms can compare dense SNP profiles to detect shared IBD segments and calculate the amount of shared DNA in centimorgans (cM). The number, size and total length of these IBD segments can then be used to infer the type of relationship.[11]

IBD segment matching algorithms have been implemented by public genetic genealogy databases where individuals can upload their DNA profiles to find their relatives.[12] These databases are populated with direct-to-consumer (DTC) profiles with hundreds of thousands of SNPs generated by private genetic testing companies.[13] There are two such databases accessible to law enforcement for forensic investigative genetic genealogy (FIGG). These are GEDmatch PRO™, the law enforcement accessible portal for the GEDmatch™ database, maintained by QIAGEN and Bode Technology, and FamilyTreeDNA, maintained by Othram Inc., both of which have over two million profiles.[14,15] However, these profiles are not accessible to law enforcement for database searching unless the owner of the profile has opted in to law enforcement searching.[16-18]

The ForenSeq® Kintelligence Kit is a targeted amplicon sequencing (TAS) panel developed by Verogen Inc. (a QIAGEN subsidiary) to allow in-house extended kinship analysis of forensic samples.[19,20] It

targets 10,230 SNPs across the human genome: 106 X chromosome SNPs, 85 Y chromosome SNPs, 56 ancestry-informative SNPs, 24 phenotype-informative SNPs, 94 identity-informative SNPs (iiSNPs) and 9,867 kinship-informative SNPs (kiSNPs). Kintelligence profiles are compatible with upload to GEDmatch PRO™ for IBD segment matching with either the one-to-one comparison to known profiles or using the one-to-many database searching function.[17] These tools have a windowed kinship algorithm that locates shared IBD segments using modified PC-AiR and PC-Relate tools that calculate kinship coefficients for the segments rather than identifying stretches of identical SNP alleles on homologous chromosomes.[21] The Kintelligence Kit has been validated and employed for casework by a number of law enforcement, academic and private laboratories.[19,22-26]

Current FIGG guidelines require that once a putative identification has been made by law enforcement, additional genetic and non-genetic testing is required to confirm the results.[27] With extended kinship inference, LR calculations could provide statistical support for such identifications, especially when suitable direct or close family reference samples are unavailable. This study sampled two pedigrees to investigate the application of STRs and SNPs for extended kinship inference using LRs and IBD segment matching. This paper presents part one of the study, which focuses on assessing the power of medium-density SNP profiling for extending the capabilities of kinship inference beyond the use of STRs.

2. Methods

2.1. Ethics Approval and Sample Procurement

Ethics approval for this research was granted by the University of Technology (UTS) Human Research Ethics Committee (HREC) (UTS HREC NO. ETH21-5821). Volunteers from Pedigree Group 1 (Individuals 1 – 12) provided self-administered buccal swabs with informed consent for DNA profiling with the Kintelligence Kit. Profiles were subsequently uploaded to GEDmatch PRO™ (Figure 1). Volunteers from Pedigree Group 2 (Individuals A – L) provided the kit numbers of pre-existing DTC kits uploaded to GEDmatch™ with informed consent to match against Kintelligence profiles uploaded to GEDmatch PRO™ (Figure 2; Table 1).[1] The central individual for Pedigree Group 2 (Individual A) also provided a self-administered buccal swab for sequencing with the Kintelligence Kit. All volunteers completed a form mapping their genetic relationships with the other volunteers in the study.

2.2. DNA Extraction and Quantification

[1] The use of DTC data for research was in compliance with the terms and conditions at the time the respective data were downloaded by the volunteers. All volunteers were notified at the conclusion of the study so they could opt out of law enforcement searching or remove their DNA data from GEDmatch™ to ensure that future use of their data is in compliance with their preferences and relevant terms and conditions.

DNA was extracted from the buccal swabs manually using the QIAamp DNA Investigator® Kit (QIAGEN) following the manufacturer's protocols for Omni Swabs with an elution volume of 100 µL.[28] Quantification was performed using the Quantifiler™ Trio DNA Quantification Kit (Thermo Fisher Scientific) on a QuantStudio™ 5 Real-Time PCR System (Thermo Fisher Scientific) following the manufacturer's protocols.[29,30] A degradation index (DI) was calculated for each sample as the ratio of concentrations of the small autosomal (SA) target and the large autosomal (LA) target.

2.3. STR Profiling

STR profiles were generated for Pedigree Group 1 samples using the GlobalFiler™ PCR Amplification Kit (Thermo Fisher Scientific).[31] The samples were amplified for 29 cycles on the Veriti™ 96-Well Fast Thermal Cycler (Thermo Fisher Scientific). The LA target was used to dilute the extracted DNA to the required input amount of 1.0 ng. Capillary electrophoresis was performed on the 3500xL Genetic Analyser (Thermo Fisher Scientific) following the manufacturer's protocols.[32] The electropherograms were analysed with GeneMapper™ ID-X v1.6 with an analytical threshold of 225 relative fluorescence units (RFU) and homozygous threshold of 1000 RFU.[33]

2.4. Library Preparation and Sequencing

Libraries were prepared manually with the Kintelligence Kit following a modified protocol.[25,34] DNA input was calculated using the LA target concentration if the DI was greater than 1 to avoid overdiluting degraded samples, otherwise the samples were diluted using the SA target concentration. Library preparation was performed in batches of 12 libraries including a positive control (NA24385; provided with the Kintelligence Kit) and a negative control (nuclease-free water). The libraries were normalised to 0.75 ng/µL using the QuantiFluor® ONE dsDNA System (Promega) on the Quantus™ Fluorometer (Promega) and pooled in batches of three libraries for sequencing.[35,36] Sequencing was performed on the MiSeq FGx® Sequencing System (QIAGEN) using a standard flow cell.[37,38]

The Kintelligence Sample Reports were exported from the Universal Analysis Software (UAS) v2.5 (QIAGEN) and analysed according to optimised thresholds in a Microsoft Excel macro described by [25].[25,39] This included a total read threshold of 20 reads and relative minor allele frequency thresholds for homozygotes (0.95 – 1.00), heterozygotes (0.10 – 0.90), sequencing error (< 0.05) and ambiguous variants (with relative frequencies in ranges 0.05 – 0.10 and 0.90 – 0.95).

2.5. Extended Kinship Analysis

STR, iiSNP (included in the Kintelligence Kit), combined identity marker (STR and iiSNP) and entire Kintelligence profiles were generated for all individuals in Pedigree Group 1. Profiles were compared using kinship LR calculation and IBD segment matching on GEDmatch PRO™. A Kintelligence profile was generated for Individual A of Pedigree Group 2. This profile was compared to DTC profiles of individuals in Pedigree Group 2 using IBD segment matching on GEDmatch PRO™.

2.5.1. LR Calculations

To calculate LRs, two alternative hypotheses were tested for each pair-wise combination of individuals in Pedigree Group 1. For unrelated pairs of individuals (19 pairs), the first hypothesis (H_1) proposed a relationship ranging from first (parent/offspring or full sibling) to eighth degree relationships and the alternative hypothesis (H_2) proposed that the individuals were unrelated. For the genetically related pairs of individuals, H_1 proposed the true relationship of the individuals and H_2 proposed that they were unrelated. For example:

H_1: Individual 1 and Individual 2 are parent/offspring.

H_2: Individual 1 and Individual 2 are unrelated members of the European population.

For each pair-wise genotype comparison, G_xG_y, LRs were calculated using the Kinship module in DBLR™ v1.3 (STRmix™)[40] according to:

$$\mathrm{LR} = \frac{\mathrm{P}(G_xG_y \mid H_1)}{\mathrm{P}(G_xG_y \mid H_2)}$$

DBLR™ does not report $\log_{10}$ LRs that exceed 300 (LR > 10^{300}) and return the $\log_{10}$ LR as infinite (∞).[40] When the relationship is impossible (LR = 0), the $\log_{10}$ LR results are given as negative infinity ($-\infty$).[40] The verbal scale for statistical support provided by $\log_{10}$ LR results is provided in Table 2.

As all individuals in this study self-identified as Australian with European ancestry, population allele frequency data for each genetic marker panel was obtained from the sources in Table 3. A Kintelligence Kit linkage map was constructed using the Map Interpolator of the Rutger's Map v.3 with sex-averaged cM positions and imported into DBLR™ to account for linkage.[41]

2.5.2. IBD Segment Matching

The Kintelligence profiles were manually edited on the UAS to be consistent with the genotypes generated with the additional thresholds in the Microsoft Excel macro. GEDmatch PRO Reports were exported from the UAS and uploaded to GEDmatch PRO™ as laboratory validation samples. Database searching was conducted using the One-to-Many Kinship tool and matches were assessed in the high confidence and expanded match lists (Table 4). For samples that were not returned in either match list, the One-to-One Kintelligence DNA Comparison tool was used to directly compare the pair of samples with their unique kit numbers.

The estimated shared cM values, average IBD segment length, number of IBD segments and number of overlapping SNPs between kits were analysed. The theoretical values were estimated using the equations of [42], modified for total length of DNA in cM, rather than Mb (Supplementary Information S1).[42] This determined the expected values for shared cM, average IBD segment length and number of IBD segments for all relationship levels tested in this study (Table 5). It should be recognised, however, that

there is not a perfect correlation between cM and Mb and so the expected cM values are only approximate.

3. Results

3.1. DNA Profiles

Table 6 outlines the call rates for the different profiles obtained for Pedigree Group 1. The average call rates exceeded 98% (STRs: 98.4 ± 1.2; iiSNPs: 98.3 ± 0.6; combined identity markers: 98.7 ± 0.5; Kintelligence: 98.4 ± 0.7). The Kintelligence profile for Individual A of Pedigree Group 2 had a call rate of 99.4%.

3.2. LR Calculations with Identity-Informative Genetic Markers (Pedigree Group 1)

3.2.1. STR Profiles

Figure 3 shows the $\log_{10}$ LRs for the true related and unrelated pairwise comparisons in Pedigree Group 1. For the first degree relationships, the parent/offspring $\log_{10}$ LRs ranged from 6.0 to 9.1 and the full sibling $\log_{10}$ LRs from 2.6 to 8.5, all providing very strong statistical support for the true relationship (H_1) as opposed to the pairs being unrelated (H_2). Log_{10} LRs for the second degree relationships ranged from 0.3 to 4.4, providing limited statistical support H_1 for an avuncular relationship (Individual 1 and Individual 4, $\log_{10}$ LR = 0.3) and grandparent/grandchild (Individual 10 and Individual 11, $\log_{10}$ LR = 0.9).

The $\log_{10}$ LRs generated for the unrelated pairs provided strong statistical support for the pairs being unrelated as opposed to being first or second degree relatives. The range of true full sibling $\log_{10}$ LRs was 2.5 to 8.5 (Figure 3). There was a false positive in the unrelated pairs with H_1 proposing a full sibling relationship (Individual 1 and Individual 12, $\log_{10}$ LR = 2.1), indicating it was approximately 120 times more likely to observe the STR profiles if the individuals were full siblings than if they were unrelated. Parent/offspring $\log_{10}$ LRs for the unrelated pairs were unable to be calculated because $P(G_xG_y \mid H_1) = 0$ as there were no alleles in common for at least one STR, implying LR = 0 ($\log_{10}$ LR = $-\infty$) and therefore, the relationship is impossible.

For third degree relationships and beyond, the $\log_{10}$ LRs were insufficient for inferring relatedness. While some of the third degree relationships generated moderate statistical support for relatedness, the majority of $\log_{10}$ LRs fell below 1 and some even below 0. The average $\log_{10}$ LR generated for the pairs of true third to eighth degree relatives was 0.21 ± 0.08 and uninformative for determining if the pair were more likely to be related or unrelated. The $\log_{10}$ LRs calculated for the unrelated pairs proposing third to eighth degree relationships were in similar ranges to the truly related pairs at these degrees, with an average $\log_{10}$ LR of -0.06 ± 0.02, and were unable to be differentiated from the results produced for pairs of true relatives. Of these unrelated pairs testing third to eighth degree relationships, 44% produced $\log_{10}$ LRs above 0.

3.2.2. iiSNP Profiles

The $\log_{10}$ LRs calculated with the 94 iiSNPs were less variable by degree than STRs for the true related pairs. However, the unrelated pairs had more variable $\log_{10}$ LRs that were typically higher than the $\log_{10}$ LRs calculated with the corresponding STR profiles. The $\log_{10}$ LRs were informative for true first and second degree relationships, with $\log_{10}$ LRs ranging from 4.6 to 9.4 for parent/offspring pairs, 4.8 to 7.2 for full sibling pairs and 1.0 to 1.9 for second degree relationships (Figure 4). One parent/offspring pair (Individual 3 and Individual 5) resulted in a $\log_{10}$ LR of $-\infty$ due to no alleles being in common at the iiSNP rs1493232 (Individual 3 typed CC from 27 reads; Individual 5 typed AA from 58 reads).

The unrelated pair observed as a false positive in the full sibling test with STRs (Individual 1 and Individual 12) also appeared as a false positive with iiSNP profiles, producing a higher $\log_{10}$ LR of 2.3. This provided moderate support for H_1, that these individuals were more likely related as full siblings than unrelated. There was an overlap in the $\log_{10}$ LRs calculated for true and false second degree relationships, with the $\log_{10}$ LRs for unrelated pairs ranging from -2.5 to 1.5.

As observed with the STR profiles, the $\log_{10}$ LRs did not provide statistical support for relatedness for third degree relationships and beyond, with an average $\log_{10}$ LR of 0.04 ± 0.04. This was not distinguishable from the $\log_{10}$ LRs calculated for the unrelated pairs, also producing an average $\log_{10}$ LR of 0.03 ± 0.02 across the third to eighth degree relationships.

3.2.3. Combined Identity Markers

The combined identity marker $\log_{10}$ LRs were calculated for both STRs (excluding SE33) and iiSNPs. This resulted in an increased statistical power compared to the individual panels and extended the informativeness of the kinship $\log_{10}$ LRs to include second degree relationships. The $\log_{10}$ LRs ranged from 11.0 to 18.5 for true parent/offspring pairs, 7.8 to 15.7 for full siblings and 1.4 to 6.0 for second degree relationships (Figure 5). Due to no typed allele in common at rs1493232, as previously discussed in the iiSNP profiles, the $\log_{10}$ LR was $-\infty$ for a parent/offspring relationship (Individual 3 and Individual 5).

The false positive observed in the unrelated STR and iiSNP profile tests for full siblings provided strong statistical support for H_1 (Individual 1 and Individual 12, $\log_{10}$ LR = 4.7), inferring it is 23,410 times more likely to observe the combined STR/iiSNP profiles if the individuals were full siblings than if they were unrelated. This was lower than the smallest $\log_{10}$ LR observed for a true full sibling relationship ($\log_{10}$ LR = 7.8). Of the ten true third degree relationships tested, half of the calculated $\log_{10}$ LRs provided only limited statistical support for H_1, while two $\log_{10}$ LRs provided limited statistical support for H_2. The average $\log_{10}$ LRs for testing third to eighth degree relationships was 0.25 ± 0.08 for the true related pairs and -0.03 ± 0.03 for the unrelated pairs.

3.3. Extended Kinship Analysis Between Kintelligence Profiles (Pedigree Group 1)

3.3.1. LR Calculations

When calculating $\log_{10}$ LRs using 10,039 autosomal SNPs available in the Kintelligence Kit, there was no overlap in $\log_{10}$ LRs for the related and unrelated pairs (Figure 6; Table 7). All pairwise comparisons with H_1 proposing a parent/offspring relationship produced $\log_{10}$ LR = -∞. Across the four pairs of true parent/offspring pairs, a total of 58 loci had no alleles in common and 97% were kiSNPs. All of these SNPs were typed homozygous in both profiles and the number of loci without a common allele ranged from six (Individual 6 and Individual 9) to 30 (individual 3 and Individual 5). The $\log_{10}$ LRs for the true full sibling relationships exceeded 300 (LR > 10^{300}), resulting in DBLR™ representing the $\log_{10}$ LR as infinite. For unrelated pairs tested with the full sibling hypothesis, the average $\log_{10}$ LR was –58.6 ± 0.5, providing strong statistical support for H_2 (Table 7).

The smallest $\log_{10}$ LRs produced for second to fifth degree relationships was 10.32 for a fifth degree relationship (Individual 3 and Individual 10). Where H_1 tested second to fifth degree relationships, the $\log_{10}$ LRs produced for true relationships were distinguishable from those produced when testing the unrelated pairs (Table 7). Of the related pairs spanning sixth to eighth degree relationships, 63% produced $\log_{10}$ LRs exceeding 4.0, providing strong statistical support for H_1. These values were highly variable, with some related pairs providing limited probative value: Individual 2 and Individual 11 with a $\log_{10}$ LR of 1.9 (eighth degree relatives), Individual 5 and Individual 10 with a $\log_{10}$ LR of 1.3 (sixth degree relatives) and Individual 5 and Individual 11 with a $\log_{10}$ LR of 0.05 (eighth degree relatives).

The $\log_{10}$ LRs produced for the unrelated pairs did not provide support for H_1 until proposing seventh degree relationships, with 16% of seventh degree and 42% of eighth degree relationship tests exceeding a $\log_{10}$ LR of 0 (Figure 6). These tests were uninformative, producing an average $\log_{10}$ LR of 0.08 ± 0.02. There was an outlier that was an exception to this (Individual 10 and Individual 12, $\log_{10}$ LR = 1.1), providing limited statistical support where H_1 proposed an eighth degree relationship. The false positive observed for the STR and iiSNP tests (Individual 1 and Individual 12) produced a $\log_{10}$ LR of -60.0 when evaluating the Kintelligence SNP profiles, showing that STRs and iiSNPs on their own are not always reliable for first degree relationship testing.

3.3.2. IBD Segment Matching

The database search using the One-to-Many Kinship tool detected 81% of the true relationships (n = 47) between the individuals in Pedigree Group 1 in either the high confidence or expanded match lists. The majority of these were fifth degree relationships or closer, of which 95% of possible first to fifth degree relationships were detected. A fourth degree relationship (Individual 2 and Individual 8) and a fifth degree relationship (Individual 4 and Individual 9) fell below the database searching thresholds (Table 4). Of the eight relationships greater than the fifth degree, only a sixth degree relationship (Individual 2 and Individual 10) was detected in the database search. The unrelated pairs in Pedigree Group 1 did not appear in the high confidence or expanded match lists as possible genetic relatives.

The undetected related pairs were directly compared in the One-to-One Kintelligence DNA Comparison tool and the fourth to seventh degree relationships could be distinguished from the true unrelated pairs. However, the eighth degree relationships (n = 3) could not be distinguished from the 19 unrelated pairs tested based on the estimated shared cM values and the number and length of IBD segments detected. Between all Kintelligence profiles compared, there was an average overlap of 9,647 ± 43 SNPs.

The estimated shared cM decreased with increased genetic distance between the pairs, with 89% of the unrelated pairs having an estimated shared cM of 0 (Figure 7). The two unrelated pairs with a single detected IBD segment were Individual 9 and Individual 11 (32 cM) and Individual 1 and Individual 12 (25 cM). While full sibling relationships were expected to share approximately 3,300 cM in total, the three full sibling relationships tested produced much lower estimated shared cM values (2,295 – 2,569 cM). Full sibling and parent/offspring relationships, both first degree relationships, were able to be differentiated as the parent/offspring estimated shared cM values were closer to the expected shared cM (3,033 – 3,194 cM). The difference between the observed and expected total shared cM values decreased with increasing genetic distance between the samples. For relationship degrees beyond third order, the expected total shared cM values were within the ranges of the observed estimated shared cM values.

The types of second degree relationships (avuncular and grandparent/grandchild) could not be differentiated from the estimated shared cM value; however, they could be distinguished by examining the number of IBD segments and their average length (Figure 8, Figure 9). The avuncular relationships had an average of 28.3 IBD segments with an overall average length of 51.4 cM. This is a greater number of IBD segments than detected in the grandparent/grandchild relationship (Individual 10 and Individual 11) with 19 IBD segments detected. However, the IBD segments were longer for the grandparent/grandchild results with an average length of 74.9 cM. While these observed values were smaller and larger than the expected values for average length and number of IBD segments, respectively, they were consistent with regard to grandparent/grandchild relationships having fewer IBD segments of greater length (Table 5).

The average length of the IBD segments detected by the GEDmatch PRO™ algorithms decreased with increased genetic distance (Figure 8). For the pairs of seventh and eighth degree relatives, and the unrelated pairs where only one IBD segment was detected, the length of that segment was recorded. The observed average length was greater than the expected average length for all relationships, with the exception of an eighth degree relationship (Individual 5 and Individual 11) where no IBD segments were detected. The largest difference between the observed and expected values was for parent/offspring relationships, with the average length for the four tested pairs ranging from 95 to 112 cM, compared to an expected average length of 61 cM.

The number of IBD segments detected between all pairs is shown in Figure 9, where there was low variability observed within the relationship degrees. The IBD segment count decreased with increased

genetic distance between the individuals. The observed number of IBD segments was lower than the expected number, but this difference decreased with increased genetic distance. The largest difference between the observed and expected IBD segment count was for the full siblings, where the results ranged from 29 to 38 IBD segments as opposed to the expected 86 IBD segments.

3.4. Extended Kinship Analysis Between Kintelligence and DTC Kits (Pedigree Group 2)

The individuals in Group 2 had up to four DTC kits each that were available for comparison to the Kintelligence profile for Individual A. All DTC kits for the known relatives of Individual A were detected in the database search on GEDmatch PRO™ in the high confidence match list. There was little variability in the results produced between the comparison of the Kintelligence profile and any of the DTC kits. Figure 10 shows the estimated shared cM by individual for each DTC kit and similar results were obtained for the parent/offspring, full sibling, second degree relatives and the unrelated pairs when compared to the Kintelligence results for Pedigree Group 1. The FamilyTreeDNA profiles produced slightly lower estimated shared cM values than Ancestry for 71% of individuals with both kits (n = 7); there was no variation observed for DTC kits of Individual A (self; 4 kits) and Individual B (parent/offspring; 4 kits). As observed in the Pedigree Group 1 study, the estimated shared cM values were lower than the expected total shared cM values calculated in Table 4.

There was little difference amongst DTC kits in the average length and number of IBD segments identified by the GEDmatch PRO™ algorithm (Figure 11, Figure 12). As observed with Pedigree Group 1, the average length of the IBD segments were higher and the number of IBD segments were lower than the expected values calculated (Table 4). All four DTC kits for Individual A had an average IBD segment length of 151.9 cM (Figure 11). The difference in IBD segment lengths across DTC kits for other individuals ranged from 1.5 cM (Individual G) to 8.6 cM (Individual C). The number of IBD segments detected for matches varied by only one segment between the DTC kits for an individual, showing low variability by DTC kit type (Figure 12). There were three individuals whose DTC kits varied by two IBD segments; the Living DNA profiles for Individual B and Individual C (parent/offspring) and the Ancestry and FamilyTreeDNA profiles for Individual G (full sibling).

There was no variability in the comparisons with the DTC kits for Individual A except for the number of overlapping SNPs; the 23andMe profile had 7,879 SNPs available for comparison with the Kintelligence profile, whereas the other three profiles had a minimum of 9,700 SNPs. Excluding this outlier (the 23andMe kit), the average number of SNPs overlapping between the Individual A Kintelligence profile and the 24 DTC kits was 9687 ± 13 SNPs.

4. Discussion

The Kintelligence Kit provides an in-house solution for forensic laboratories with massively parallel sequencing technology to expand their kinship inference capabilities. Whereas STR profiles have been historically limited to short-range kinship analysis for first degree relationships, sequencing several

thousand SNPs enables medium-range kinship analysis to reliably distinguish between related and unrelated pairs out to the fifth degree (e.g. second cousins) for both IBD segment matching and LR calculations. The medium-density genotypes generated with the Kintelligence Kit can be analysed using the traditional LR approach, providing statistical support for relatedness in accordance with international recommendations.[43] Alternatively, IBD segment matching can be applied using a windowed kinship algorithm to facilitate FIGG by estimating shared haplotype distributions.[17,21] Furthermore, this analytical method allows profiles derived from forensic samples to be compared to DTC profiles uploaded by members of the public that have given permission for them to be searched by law enforcement.

The current guidelines for kinship analysis recommend calculating LRs for first degree relatives (i.e. parent, offspring and full sibling) using routine STR profiling.[1,2] The basis for this approach is that first degree relatives will share more alleles than unrelated individuals.[5,8] Searching law enforcement databases for indirect matches allows testing of the propositions that a pair of DNA donors are parent/offspring or unrelated, or that the pair are full siblings or unrelated.[1,5]

Several studies have recommended prioritising parent/offspring familial reference samples over full siblings as the direct inheritance of chromosomes ensures at least one allele in common at every locus.[44,45] This recommendation is supported by the STR analysis in this study, which produced strong $\log_{10}$ LRs supporting relatedness for the parent/offspring pairs. However, the false positive results observed when testing a full sibling relationship between unrelated individuals demonstrated that STRs and iiSNPs alone are not always reliable for this relationship. Kinship analysis using the 94 iiSNPs performed similarly to STRs, with the resulting kinship LRs only slightly more powerful with the iiSNPs. This is consistent with a study that assessed the discriminating power of STRs and iiSNPs for direct comparisons of casework-type samples, where the average $\log_{10}$ LR was 23 for STRs and 38 for the 94 iiSNPs.[46]

For both STRs and iiSNPs, there were too few markers to differentiate more distant genetic relatives from unrelated individuals in this study. The unrelated pairs typically produced higher $\log_{10}$ LRs when comparing iiSNP profiles as opposed to STR profiles; this was not unexpected due to the biallelic nature of the 94 iiSNPs in this study, resulting in only three possible genotypes at a locus and increasing the probability of alleles being identical by chance. [47] proposed combining the LRs of STRs and SNPs to improve kinship probabilities for partial profiles obtained in an ancient DNA case study.[47] Similarly, this study showed a substantial increase in the $\log_{10}$ LRs when combining identity markers. The range of kinship analysis could be expanded to include second degree relationships when applying the combined identity marker panel; however, this still produced a false positive result for the full sibling tests between an unrelated pair. This approach can be applied to short-range kinship analysis to test first and second degree relationships and improve the statistical support for relatedness when comparing partial profiles.

When LRs were calculated with 10,039 autosomal SNPs targeted by the Kintelligence Kit, powerful statistical support allowed for true full siblings to fifth degree relatives (e.g. second cousins) to be

differentiated from unrelated individuals. Log_{10} LRs could not be calculated for the parent/offspring relationships with the Kintelligence genotypes because at least one locus had no shared alleles between the profiles. This is because DBLR™ applies a binary approach to test parent/offspring relationships, requiring a shared allele at every locus. Genotyping errors, resulting in allele dropout, are expected with a panel of this size, with up to 30 SNPs (0.3% of the profile) showing no shared alleles between true parent/offspring pairs in this study.[48] It may be possible to apply a probabilistic model that accounts for genotyping error when calculating LRs, such as that published by [48].[48]

The majority of the relationships spanning full sibling to seventh degree (93%) produced log_{10} LRs that exceeded the extremely strong support threshold (log_{10} LR = 6) and 98% exceeding the very strong support threshold (log_{10} LR = 5).[43] These results are concordant with the findings of simulations of medium-density SNP panels to test kinship LR calculations.[3,9,49] Gettings *et al.* (2024) simulated Kintelligence data for first to fifth degree relationships and produced similar results, calculating median log_{10} LRs of 1300 for first degree, 300 for second degree, 120 for third degree, 50 for fourth degree and 20 for fifth degree relationships. Notably, 95% of fifth degree relationship log_{10} LRs exceeded 4, providing strong support for relatedness.[49]

A study simulating 6,600 SNPs for second to sixth degree relatives with direct lineage or one or two common ancestors concluded that while denser marker sets can produce more powerful LRs, they also increase the number of false positive relationships.[3] In this study, the unrelated pairwise comparisons with the Kintelligence profiles did not produce log_{10} LRs that falsely supported a relationship, except for one test hypothesising a seventh degree relationship. However, this provided only limited support for relatedness and could be differentiated from the log_{10} LRs generated for the true seventh degree relationships. Furthermore, the false positive result supporting a full sibling relationship between an unrelated pair with STRs and iiSNPs was eliminated when using Kintelligence genotypes.

IBD segment matching algorithms have become increasingly common tools for kinship analysis in forensics, particularly with the use of FIGG for long-range kinship testing.[12] The windowed kinship algorithm on GEDmatch PRO™ infers IBD segments by first assessing stretches of SNPs with at least one allele in common, followed by a second pass to identify SNPs within those segments that share both alleles.[21] While the algorithm was originally developed and validated for first to fourth degree relationships, this study found it was also capable of detecting fifth degree relatives in database searches, as well as facilitating direct comparisons to known profiles for sixth and seventh degree relatives. Furthermore, the number and average length of the detected IBD segments can be taken into account to differentiate relationships within the same degree, as demonstrated in this study with second degree relationships (avuncular and grandparent/grandchild). These variations are attributed to the number of meioses separating the queried individuals, with additional meiosis leading to smaller IBD segments being shared between them.[11,42]

When developing the windowed kinship algorithm, it was concluded that the estimated shared cM values produced by this algorithm would be, on average, 7% lower than those produced by the GEDmatch™ algorithm, which was developed for high-density SNP data analysis.[21] In this study, the estimated shared cM values were, on average, 16% lower than the expected values. It was not unexpected that the observed shared cM, average length of IBD segments and number of IBD segments were different to the expected values. This is because the calculations were based on the entire genome of approximately three billion nucleotides, whereas the Kintelligence Kit targets only 10,230 SNPs or approximately 0.00034% of the human genome. While the kiSNPs were selected to target linked SNPs along the chromosomes that would correlate with IBD segments, this is still an estimation, and the available data is more limited compared to microarray or whole genome sequencing (WGS) approaches. Microarrays, which are commonly employed by DTC companies for their genetic testing, target anywhere from 500,000 to 2,500,000 nucleotides, while WGS aims to sequence the entire human genome of three billion nucleotides.[12]

The theoretical values calculated with the modified equations by Speed & Balding (2015) were lower than those published on the Shared cM Project 4.0 Tool v4 with the exception of full siblings.[42,50,51] The values published on this tool are based on data uploaded to the website by the public as opposed to calculations with sex-averaged recombination numbers per meiosis. The observed shared cM values for the full sibling relationships in this study were closer to those published on the Shared cM Project 4.0 than the expected shared cM values calculated. It is also important to note that there is not a perfect correlation between cM and Mb and so the expected cM values calculated by modifications to the equations of Speed & Balding (2015; in Mb) are only approximations but should be reasonable approximations for average values.[42] The expected number of IBD segments, however, should not depend on the unit chosen for IBD segment length.

GEDmatch PRO™ provides expected cM ranges and the average cM in the Generation Chart as part of the One-to-Many Kinship tool, which are generally lower than the values published on the Shared cM Project.[17, 52, 53] However, the Generation Chart groups first degree relationships together, giving the expected range of 2,328 to 3,065 cM with an average of 2,697 cM. The Shared cM Project separates parent/offspring (range of 2,376 – 3,720 cM, average of 3,485 cM) and full sibling (range of 1,613 – 3,488 cM, average of 2,613 cM) values and, while there is overlap in the expected ranges for these relationships, the parent/offspring relationships are typically several hundred cM higher.[17, 52, 53] The estimated shared cM observed in this study for the first degree relationships fell within the broad ranges observed by the Shared cM Project.[52, 53] However, some results did not fall within the ranges provided by the Generation Chart on GEDmatch PRO™.[17] This study supports differentiation of the parent/offspring and full sibling expected shared cM in the Generation Chart for classification of first degree relationships.

The results obtained when Kintelligence profiles were compared to either other Kintelligence profiles (Pedigree Group 1) or DTC kits (Pedigree Group 2) were consistent, with no observable impact from the company providing the DTC testing service. Other studies assessing Kintelligence and microarray results have observed minimal variation in the estimated shared cM across different sequencing methods, confirming the Kintelligence Kit is reliable for detecting and classifying first to fourth degree relationships.[52,53] This lack of variation is due to the Kintelligence Kit being specifically designed to target SNPs common to various DTC kits and those used by the existing GEDmatch™ algorithm for high-density SNP data.[19,21] This allows for public databases and profiles uploaded by the public to be leveraged for extended kinship analysis, eliminating the need for the creation of a new law enforcement database containing reference SNP profiles.

[9] observed that the IBD segment matching approach for kinship analysis was more effective at excluding unrelated individuals, while LR calculations were able to classify related individuals more accurately in simulations using high-density SNP profiles.[9] However, in this study, the expanded match list generated following a database search on GEDmatch PRO™ often included individuals who were not known genetic relatives of the volunteers. Previous studies have concluded that, while absolute determination of a relationship is not possible, it is crucial to assess which relationship is most probable.[54] While the LR approach requires testing multiple sets of hypotheses to infer the most likely relationship between individuals, the degree of relatedness can be inferred from the total amount of DNA shared, the average shared IBD segment length and the number of shared IBD segments using the IBD segment matching approach.

The Kintelligence Kit has been successfully used to resolve both criminal and unidentified human remains cases around the world, relying on genetic genealogy databases and IBD segment matching algorithms. However, FIGG requires confirmatory testing of the proposed identification using either a direct match with a PoI or partial matches and kinship LR calculations with STR profiles from first degree relatives of the PoI.[55,56] This study demonstrates that applying the LR approach to Kintelligence profiles is a suitable alternative when these preferred family reference samples are unavailable, allowing for confirmation of identity using more distant relatives with statistical support. Ideally, it would be preferable to provide support for genetic relationships on more than one lineage of a pedigree. However, forensic casework samples are often of poor quality and further evaluation is required to determine how extended kinship analysis should be approached with partial profiles. Part two of this study, published separately, assesses how genotyping error and information loss impact the ability to calculate LRs and detect IBD segments.[59]

5. Conclusions

This study demonstrates that the Kintelligence Kit can extend kinship inference beyond short-range familial searches using STR profiles. While identity-informative markers such as STRs and iiSNPs are suitable for first degree relationships, the Kintelligence Kit can be applied for kinship analysis of first to

fifth degree relationships using kinship LR calculations and IBD segment matching. The SNP genotypes can be analysed with kinship LRs to provide statistical support for hypotheses of relatedness or uploaded to law enforcement accessible genetic genealogy databases for analysis using the windowed kinship algorithm to infer IBD segments. Part two of this study builds on the assessment of the Kintelligence Kit and various kinship analysis methods, exploring the impact of suboptimal Kintelligence profiles, locus dropout and allele dropout on kinship analysis.

6. References

1. National Institute of Justice. Familial DNA Searching: Current Approaches. National Institute of Justice; 2015;1-82. Available from: https://www.ojp.gov/library/publications/familial-dna-searching-current-approaches [Last accessed: 11/18/2024].
2. Scientific Working Group on DNA Analysis Methods. Recommendations from the SWGDAM Ad Hoc Working Group on Familial Searching. SWGDAM Communications; 2013. Available from: https://www.swgdam.org/publications [Last accessed: 01/02/2025].
3. Kling D. On the use of dense sets of SNP markers and their potential in relationship inference. Forensic Sci Int Genet 2019;39:19-31; doi: 10.1016/j.fsigen.2018.11.022.
4. Maguire CN, McCallum LA, Storey C, et al. Familial searching: A specialist forensic DNA profiling service utilising the National DNA Database® to identify unknown offenders via their relatives—The UK experience. Forensic Sci Int Genet 2014;8(1):1-9; doi: 10.1016/j.fsigen.2013.07.004.
5. Ge J, Chakraborty R, Eisenberg A, et al. Comparisons of Familial DNA Database Searching Strategies. J Forensic Sci 2011;56(6):1448-1456; doi: 10.1111/j.1556-4029.2011.01867.x.
6. Schwab AP, Luu HS, Wang J, et al. Genomic Privacy. Clin Chem 2018;64(12):1696-1703; doi: 10.1373/clinchem.2018.289512.
7. Kennett D. Using genetic genealogy databases in missing persons cases and to develop suspect leads in violent crimes. Forensic Sci Int 2019;301:107-117; doi:10.1016/j.forsciint.2019.05.016.
8. Ge J, Budowle B, Chakraborty R. Choosing Relatives for DNA Identification of Missing Persons. J Forensic Sci 2011;56:S23-S28; doi: 10.1111/j.1556-4029.2010.01631.x.
9. Kling D, Tillmar A. Forensic genealogy—A comparison of methods to infer distant relationships based on dense SNP data. Forensic Sci Int Genet 2019;42:113-124; doi: 10.1016/j.fsigen.2019.06.019.
10. Ge J, Budowle B. Forensic investigation approaches of searching relatives in DNA databases. J Forensic Sci 2021;66(2):430-443; doi: 10.1111/1556-4029.14615.
11. Morimoto C, Manabe S, Fujimoto S, et al. Discrimination of relationships with the same degree of kinship using chromosomal sharing patterns estimated from high-density SNPs. Forensic Sci Int Genet 2018;33:10-16; doi: 10.1016/j.fsigen.2017.11.010.
12. Kling D, Phillips C, Kennett D, et al. Investigative genetic genealogy: Current methods, knowledge and practice. Forensic Sci Int Genet 2021;52:102474; doi: 10.1016/j.fsigen.2021.102474.

13. Glynn CL. Bridging Disciplines to Form a New One: The Emergence of Forensic Genetic Genealogy. Genes 2022;13(8):1381,; doi: 10.3390/genes13081381.

14. GEDmatch. Find DNA Matches for Free | Analyze Your DNA | GEDmatch.; 2024. Available from: https://www.gedmatch.com/ [Last Accessed; 25/11/25/2024].

15. FamilyTreeDNA. DNA Testing for Ancestry & Genealogy | FamilyTreeDNA; 2024. Available from: https://www.familytreedna.com/ [Last Accessed; 25/11/25/2024].

16. GEDmatch. GEDmatch.com Terms of Service and Privacy Policy; 2024. Available from: https://www.gedmatch.com/terms-of-service-privacy-policy [Last Accessed; 25/11/25/2024].

17. GEDmatch PRO. GEDmatch PRO.; 2024. Available from: https://pro.gedmatch.com/ [Last accessed 11/25/2024].

18. FamilyTreeDNA. FamilyTreeDNA Law Enforcement Guide.; 2024. Available from: https://www.familytreedna.com/legal/law-enforcement-guide [Last Accessed; 25/11/25/2024].

19. Antunes J, Walichiewicz P, Forouzmand E, et al. Developmental Validation of the ForenSeq® Kintelligence Kit, MiSeq Fgx® Sequencing System and ForenSeq Universal Analysis Software. Forensic Sci Int Genet 2024;71:(103055;, doi: 10.1016/j.fsigen.2024.103055.

20. QIAGEN. Verogen ForenSeq Kintelligence Kit; 2024. Available from: https://www.qiagen.com/us/products/human-id-and-forensics/nextgeneration-sequencing/verogen-forenseq-kintelligence-kit [Last Accessed; 24 May05/24/2024].

21. Snedecor J, Fennell T, Stadick S, et al. Fast and accurate kinship estimation using sparse SNPs in relatively large database searches. Forensic Sci Int Genet 2022;61:(102769,; doi: 10.1016/j.fsigen.2022.102769.

22. Arenson J. DNA Labs International is the first accredited lab to use the ForenSeq® Kintelligence System to aid in the identification of unidentified remains. PRWeb: Deerfield Beach, FLA; 2022. Available from: https://biotechnetworks.org/2022/02/23/dna-labs-international-is-the-first-accredited-lab-to-use-the-forenseq-kintelligence-system-to-aid-in-the-identification-of-unidentified-remains/ [Last accessed: 01/10/2025].

23. Signature Science. SigSci's Center for Advanced Genomics Validating the Verogen ForenSeq® Kintelligence Workflow for Forensic Genetic Genealogy. Austin, TX; 2021. Available from: https://www.signaturescience.com/press/ngs_fgg/ [Last Accessed; 31/03/31/2022].

24. Peck MA, Koeppel AF, Gorden EM, et al. Internal Validation of the ForenSeq Kintelligence Kit for Application to Forensic Genetic Genealogy. Forensic Genom 2022;2(4): 103-114,; doi: 10.1089/forensic.2022.0014.

25. Watson JL, McNevin D, Grisedale K, et al. Operationalisation of the ForenSeq® Kintelligence Kit for Australian Unidentified and Missing Persons Casework. —Forensic Sci Int Genet 2023; 68:102972;, doi: 10.1016/j.fsigen.2023.102972.

26. Jordan S. CHI becomes nation's first public lab to earn accreditation to perform Forensic Genetic Genealogy; 2023. Available from: https://www.unthsc.edu/newsroom/story/chi-becomes-nations-first-public-lab-to-earn-accreditation-to-perform-forensic-genetic-genealogy/ [Last Accessed; 05/02/05/2024].

27. Department of Justice. Forensic Genetic Genealogical DNA Analysis and Searching. Interim Policy; 2019. Available from: https://www.justice.gov/olp/page/file/1204386/dl [Last accessed 20/11/2024].

28. QIAGEN. QIAamp® DNA Investigator Handbook.; 2020. Available from: https://www.qiagen.com/us/products/human-id-and-forensics/investigator-solutions/qiaamp-dna-investigator-kit [Last accessed 01/05/2024].

29. Thermo Fisher Scientific. Quantifiler™ HP and Trio DNA Quantification Kits User Guide [User Guide].; 2017. Available from: https://www.thermofisher.com/order/catalog/product/4482910 [Last accessed 01/05/2024].

30. Thermo Fisher Scientific. QuantStudio™ 5 Real-Time PCR Instrument (for Human Identification) [User Guide]. Life Technologies; 2017. Available from: https://www.thermofisher.com/order/catalog/product/A34322 [Last accessed: 01/05/2024].

31. Thermo Fisher Scientific. GlobalFiler PCR Amplification Kits;. 2019. Available from: https://www.thermofisher.com/order/catalog/product/4476135 [Last accessed 01/05/2024].

32. Thermo Fisher Scientific. Applied Biosystems 3500/3500xL Genetic Analyzer User Guide.; 2010. Available from https://www.thermofisher.com/order/catalog/product/4406016 [Last accessed: 01/05/2024].;

33. Thermo Fisher Scientific. GeneMapper™ ID-X Software v1.6. 2024. Available from: https://www.thermofisher.com/order/catalog/product/A39975 [Last accessed: 01/05/2024].

34. Verogen. ForenSeq Kintelligence Kit Reference Guide. ForenSeq Kintelligence Kit; 2021. Available from: https://verogen.com/products/forenseq-kintelligence-kit/ [Last accessed 02/19/2024].

35. Promega. Quantus™ Fluorometer;. 2022. Available from: https://www.promega.com.au/products/microplate-readers-fluorometers-luminometers/fluorometers/quantus-fluorometer/?catNum=E6150 [Last accessed 03/29/2024].

36. Promega. QuantiFluor® ONE dsDNA System.; 2022. Available from: https://www.promega.com.au/products/rna-analysis/dna-and-rna-quantitation/quantifluor-dsdna-system/?catNum=E2670 [Last accessed 03/29/2024].

37. Verogen. MiSeq FGx Sequencing System Reference Guide;. MiSeq FGx Sequencing System 2021. Available from: https://verogen.com/products/miseq-fgx-sequencing-system/ [Last accessed: 02/19/2024].

38. QIAGEN. Verogen MiSeq FGx Sequencing Reagents.; 2024. Available from: https://www.qiagen.com/us/products/human-id-and-forensics/nextgeneration-sequencing/verogen-miseq-fgx-sequencing-reagents [Last Accessed; 08/15/2024 August].

39. Verogen. Universal Analysis Software v2.0 and above, Reference Guide.; Universal Analysis Software 2021. Available from: https://verogen.com/products/universal-analysis-software/ [Last accessed: 02/19/2024].

40. STRmix™. DBLR™;. 2023. Available from: strmix.com/dblr/ [Last Accessed; 01/29/2024 January].

41. University R. Map Interpolator of the Rutgers Map;. 2014. Available from: http://compgen.rutgers.edu/map_interpolator.shtml [Last Accessed; 10/20/2024 October].

42. Speed D, Balding DJ. Relatedness in the post-genomic era: is it still useful? Nat Rev Genet 2015;16(1):33-44;, doi: 10.1038/nrg3821.

43. Scientific Working Group on DNA Analysis Methods. Recommendations of the SWGDAM Ad Hoc Working Group on Genotyping Resulted Reported as Likelihood Ratios;- 2016. Available from: https://www.swgdam.org/publications [Last accessed: 01/02/2025].

44. Zvénigorosky V, Sabbagh A, Gonzalez A, et al. The limitations of kinship determinations using STR data in ill-defined populations. Int J Legal Med 2020;134(6):1981-1990;; doi: 10.1007/s00414-020-02298-w.

45. Ge J, Budowle B. How many familial relationship testing results could be wrong? PLoS Genet 2020;16(8):e1008929;, doi: 10.1371/journal.pgen.1008929.

46. Watson JL, Grisedale K, Cho K, et al. Characterisation of Identity-Informative Genetic Markers in the Australian Population with European Ancestry. Forensic Sci Int Genet 2025;74:103169; doi: 10.1016/j.fsigen.2024.103169. [pre-print]. Available at SSRN 4854523 2024;

47. Pajnič IZ, Leskovar T, Črešnar M. Improving kinship probability in analysis of ancient skeletons using identity SNPs and MPS technology. Int J Legal Med 2023;137(:1007-1015;; doi: 10.1007/s00414-023-03003-3.

48. Andersen MM, Kampmann M-L, Jepsen AH, et al. Shotgun DNA sequencing for human identification: Dynamic SNP selection and likelihood ratio calculations accounting for errors. Forensic Sci Int Genet 20254;74:103146;, doi: 10.1016/j.fsigen.2024.103146.

49. Gettings KB, Tillmar A, Sturk-Andreaggi K, et al. Review of SNP assays for disaster victim identification: Cost, time, and performance information for decision-makers. J Forensic Sci 2024;69(5):1546-1557;, doi: 10.1111/1556-4029.15585.

50. Bettinger B. The shared cM project: a demonstration of the power of citizen science. J Genet Geneal 2016;8(1):38-42. Available from: https://jogg.info/wp-content/uploads/2021/09/81.006.pdf.)

51. Bettinger BT. The Shared cM Project 4.0 tool v4;- 2020. Available from: https://dnapainter.com/tools/sharedcmv4 [Last Accessed; 25/11/25/2024].

52. Hymus CM, Tay JW, Coghlan ML, et al. Tree vs tree–comparing kinship prediction outcomes for three generations of a family using the Verogen Kintelligence and Infinium Global Screening Array-24 Kit. Aust J Forensic Sci 2024;56(sup1):119-122;, doi: 10.1080/00450618.2024.2324750.

53. Daniel R, Raymond J, Sears A, et al. It's all relative: A multi-generational study using ForenSeq™ Kintelligence. Forensic Sci Int 2024;364:(112208;, doi: 10.1016/j.forsciint.2024.112208.

54. Weir BS, Anderson AD, Hepler AB. Genetic relatedness analysis: modern data and new challenges. Nat Rev Genet 2006;7(10):771-780;, doi: 10.1038/nrg1960.

55. Greytak EM, Moore C, Armentrout SL. Genetic genealogy for cold case and active investigations. Forensic Sci Int 2019;299:(103-113, doi:.10.1016/j.forsciint.2019.03.039.

56. Scientific Working Group on DNA Analysis Methods. Overview of investigative genetic genealogy. SWGDAM Communications; 2020. Available from: https://www.swgdam.org/publications [Last accessed: 01/02/2025].

57. Taylor D, Bright J-A, McGovern C, et al. Allele frequency database for GlobalFiler™ STR loci in Australian and New Zealand populations. Forensic Sci Int Genet 2017;28:(e38-e40;, doi: 10.1016/j.fsigen.2017.02.012.

58. The International Genome Sample Resource. 1000 Genomes Project;- 2021. Available from: https://www.internationalgenome.org/ [Last Accessed; 10/17/2023 October].

59. Watson JL, Grisedale K, Coakley L, et al. Extended Kinship Inference. Part 2: Evaluation of the Impact of Information Loss on Likelihood Ratios and Haplotype Matching [In Press].

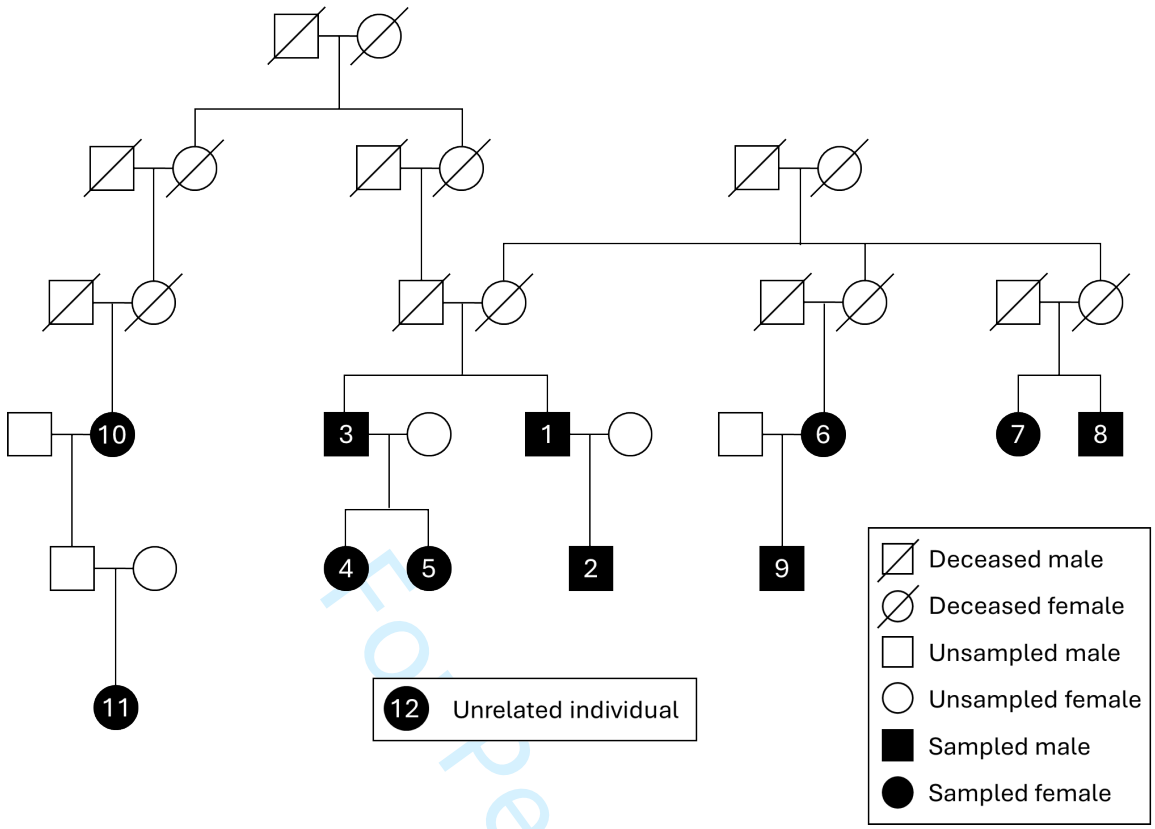

Figure 1. Pedigree Group 1, with numbers corresponding to Individuals 1 through 12. Individuals included in the study are filled in (black), with living relatives linking the family members not filled (white). Deceased relatives are crossed out. Individual 12 is unrelated to all other individuals in the group.

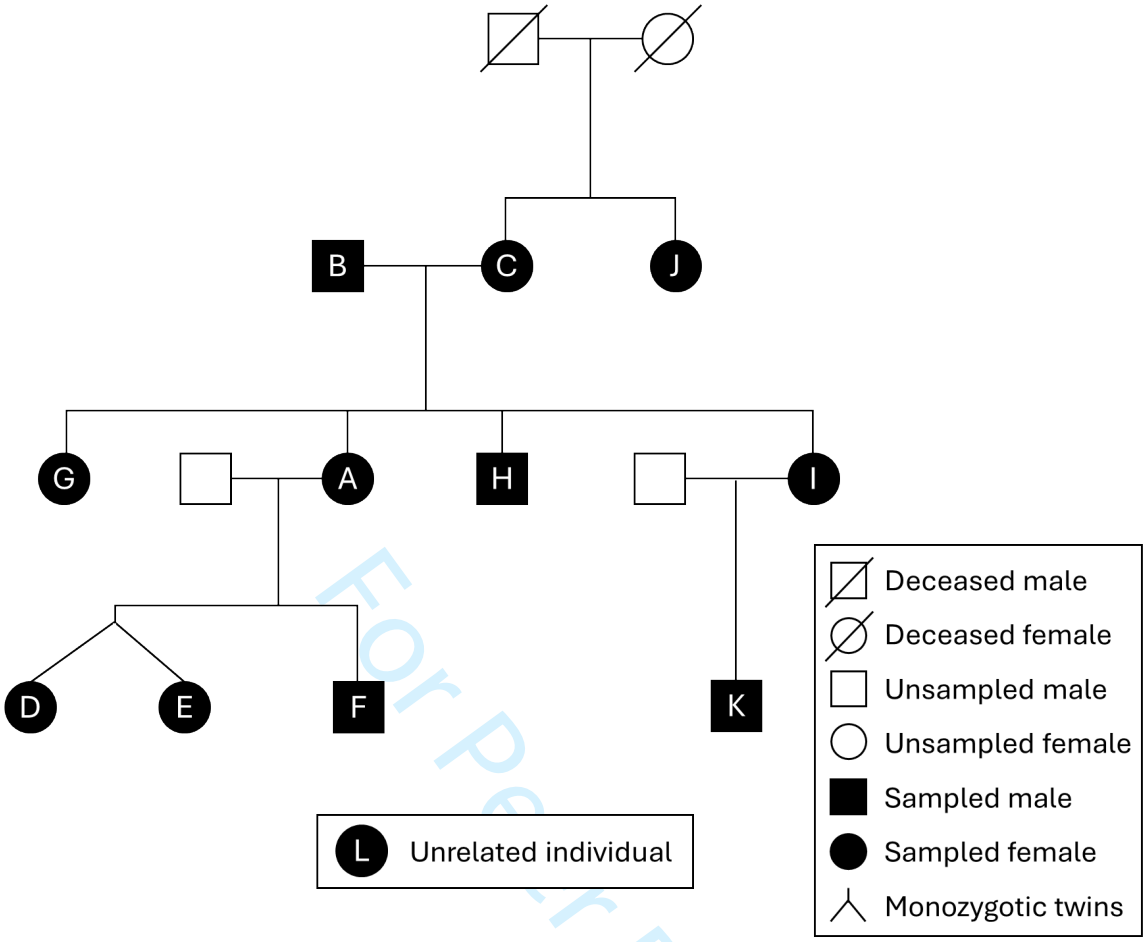

Figure 2. Pedigree Group 2, with letters corresponding to Individuals A through L. Individuals included in the study are filled in (black), with living relatives linking the family members not filled (white). Deceased relatives are crossed out. Individual L is unrelated to all other individuals in the group.

Table 1. Direct-to-consumer kits available for Individuals A through L in Group 2.

Sample ID	AncestryDNA	FamilyTreeDNA	23andMe	Living DNA
Individual A	✓	✓	✓	✓
Individual B	✓	✓	✓	✓
Individual C	✓	✓	✓	✓
Individual D	×	✓	×	×
Individual E	✓	✓	×	×
Individual F	×	✓	×	×
Individual G	✓	✓	×	×
Individual H	✓	✓	×	×
Individual I	✓	✓	×	×
Individual J	×	✓	×	×
Individual K	×	✓	×	×
Individual L	✓	×	×	×

Table 2. Verbal scale for log_{10} likelihood ratio (LR) results and the statistical support for each hypothesis.[43]

Log_{10} LR	Interpretation
≤ –5	Very strong statistical support for H_2
–4	Strong statistical support for H_2
–3	Moderately strong statistical support for H_2
–2	Moderate statistical support for H_2
–1	Limited statistical support for H_2
0	Uninformative
1	Limited statistical support for H_1
2	Moderate statistical support for H_1
3	Moderately strong statistical support for H_1
4	Strong statistical support for H_1
≥ 5	Very strong statistical support for H_1

Table 3. Population allele frequency data utilised for each marker set: short tandem repeat (STR), identity-informative single nucleotide polymorphism (iiSNP), combined identity markers (STRs and iiSNPs), and all Kintelligence SNPs.

Market Set	Panel	Markers	Population Data	Source
STRs	GlobalFiler™ PCR Amplification Kit[31]	21	Australian Caucasian	Taylor *et al.* (2017)[57]
iiSNPs	ForenSeq® Kintelligence Kit[34]	94	Australian with European Ancestry	Watson *et al.* (2024)[46]
Combined Identity Markers	GlobalFiler™ PCR Amplification Kit and ForenSeq® Kintelligence Kit[31,34]	114[a]	Australian Caucasian (STR) Australian with European Ancestry (iiSNP)	Taylor *et al.* (2017) and Watson *et al.* (2024)[46,57]
Kintelligence SNPs	ForenSeq® Kintelligence Kit[34]	10,039[b]	European	1000 Genomes[58]

[a] SE33 was excluded as linkage equilibrium tests could not be performed between SE33 and iiSNPs.[46]

[b] X and Y SNPs were excluded from the Kintelligence Kit for likelihood ratio calculations.

Table 4. Shared centimorgan (cM) thresholds applied in the GEDmatch PRO™ One-to-Many Kinship tool match lists based on the number of overlapping SNPs. For all threshold levels, the longest identical-by-descent (IBD) segment needs to be at least 30 cM.[17]

Overlapping SNPs	Shared cM Threshold	
	High Confidence Match List	Expanded Match List
9000	170	120
8000	190	140
6000	200	160

Table 5. Expected values for total shared centimorgan (cM), average identical by descent (IBD) segment length and number of IBD segments calculated for different relationships. Values were calculated with the equations in Supplementary Information S1.

Kinship Degree	Relationship	Total Shared cM	Average IBD Segment Length (cM)	Number of IBD Segments
1	Parent / offspring	3300	61.3	53.8
1	Full sibling	3300	38.6	85.6
2	Avuncular	1650	28.1	58.7
2	Grandparent / grandchild	1650	38.6	42.8
3	First cousin	825	22.1	37.3
4	First cousin once removed	412.5	18.2	22.6
5	Second cousin	206.3	15.5	13.3
6	Second cousin once removed	103.1	13.5	7.6
7	Second cousin twice removed	51.6	11.9	4.3
8	Third cousin once removed	25.8	10.7	2.4
N/A	Unrelated	0	0	0

Table 6. Call rates for the DNA profiles generated for each sample in Pedigree Group 1; calculated for short tandem repeats (STRs; 21), identity-informative SNPs (iiSNPs; 94), combined identity markers (STRs and iiSNPs; 114) and autosomal Kintelligence SNPs (10,039).

Sample ID	STRs (%)	iiSNPs (%)	Combined Identity Markers (114)	Kintelligence (%)
Individual 1	100.00	100.00	100.00	99.86
Individual 2	100.00	100.00	100.00	99.82
Individual 3	100.00	96.81	98.25	96.96
Individual 4	100.00	95.74	96.49	97.85
Individual 5	85.71	97.87	97.80	96.18
Individual 6	100.00	100.00	100.00	99.76
Individual 7	100.00	98.94	99.12	99.76
Individual 8	100.00	100.00	100.00	99.95
Individual 9	100.00	98.94	99.12	99.74
Individual 10	95.24	92.55	93.86	92.02
Individual 11	100.00	100.00	100.00	99.79
Individual 12	100.00	98.94	99.12	99.41

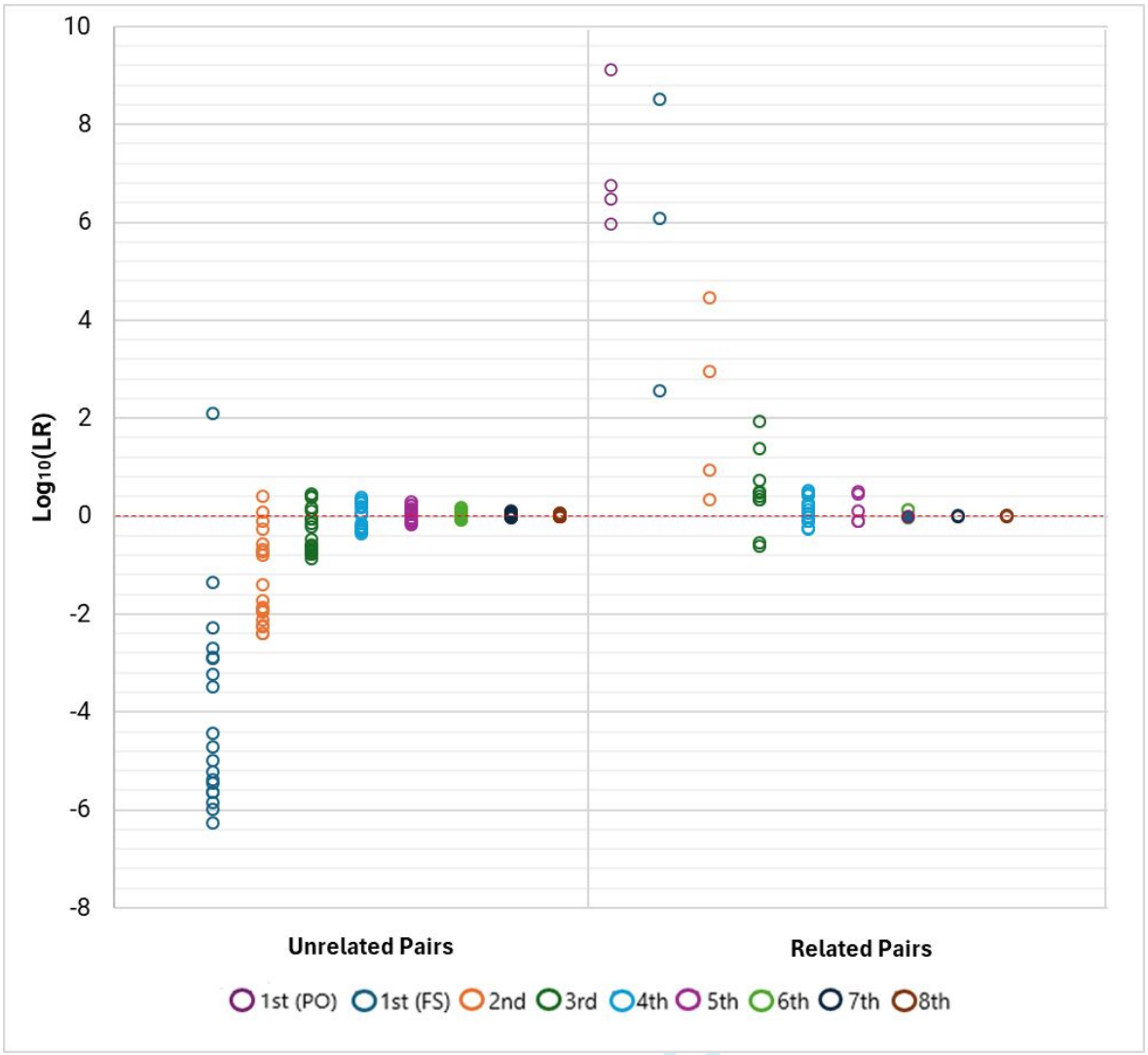

Figure 3. Log_{10} LR generated using the short tandem repeats (STRs) in the GlobalFiler™ PCR Amplification Kit for each relationship degree for true related pairs (n = 47) and all relationship degrees for unrelated pairs (n = 152) in Pedigree Group 1. The dotted red line marks uninformative results (log_{10} LR = 0). Unrelated pairs have not been plotted for parent/offspring tests ($\log_{10}$ LR = $-\infty$). PO: parent/offspring; FS: full sibling.

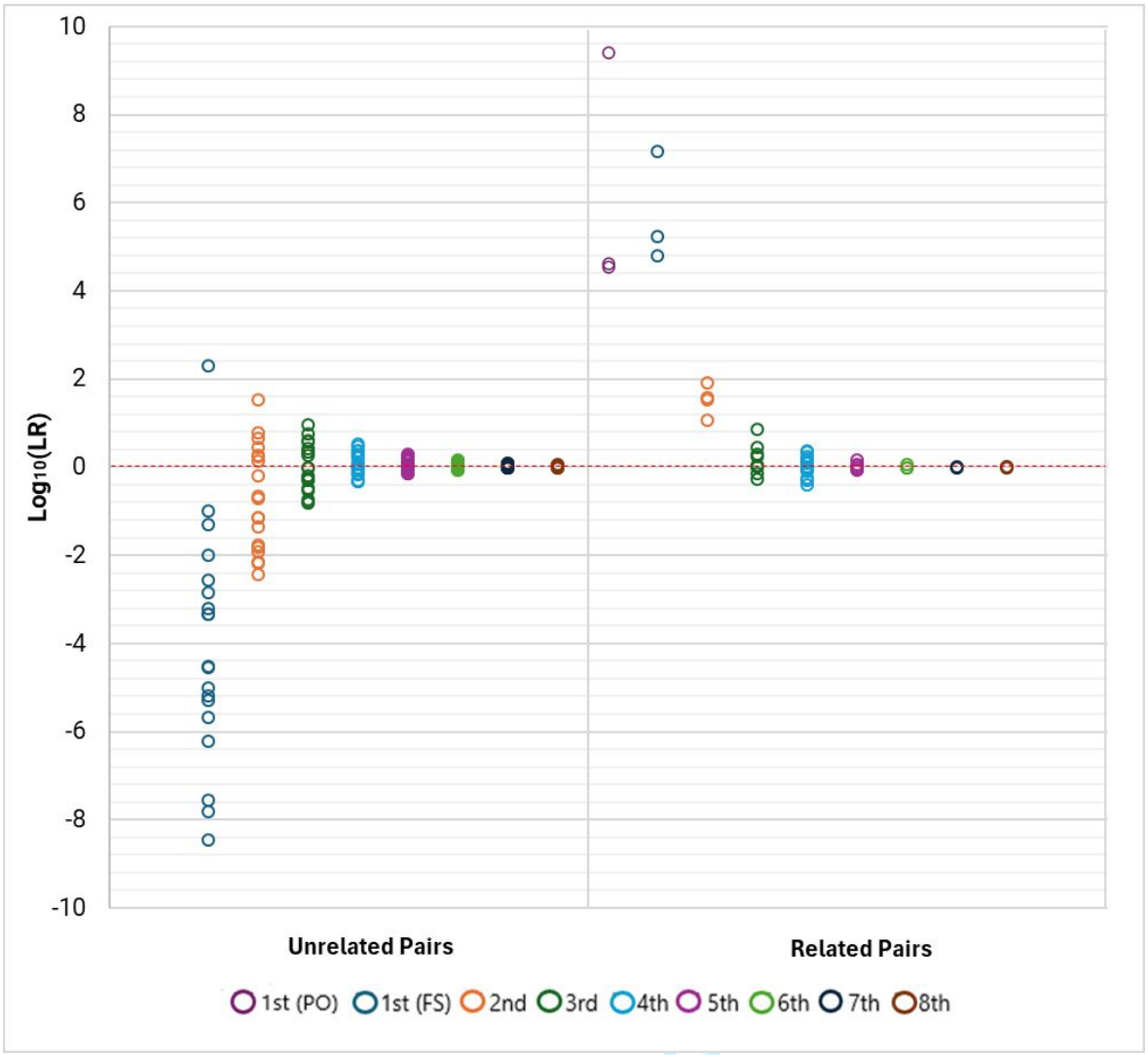

Figure 4. Log_{10} LR generated using the identity-informative single nucleotide polymorphisms (iiSNPs) in the ForenSeq® Kintelligence Kit for each relationship degree for true related pairs (n = 47) and all relationship degrees for unrelated pairs (n = 152) in Pedigree Group 1. The dotted red line marks uninformative results (log_{10} LR = 0). Unrelated pairs have not been plotted for parent/offspring tests (log_{10} LR = -∞). PO: parent/offspring; FS: full sibling.

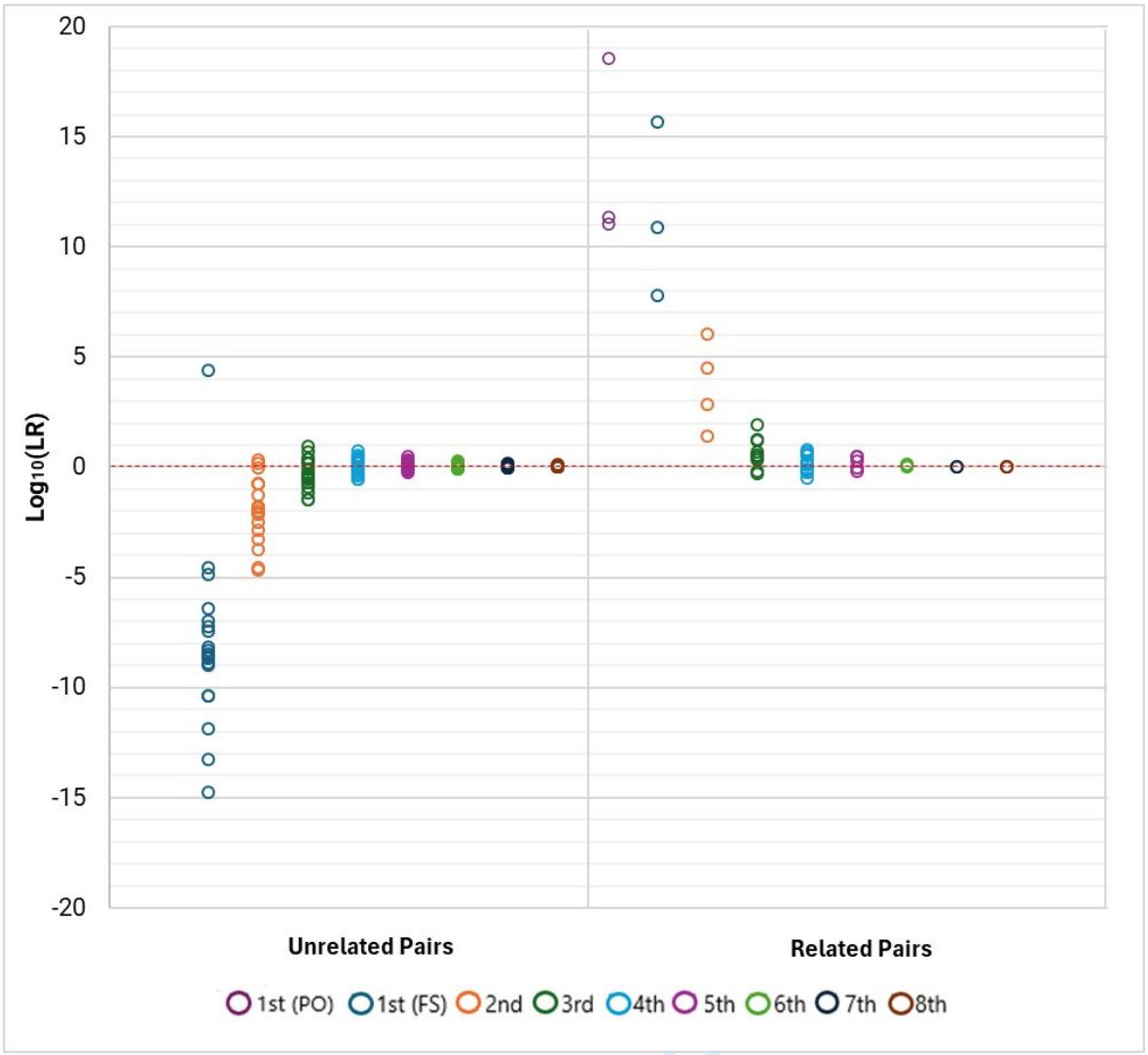

Figure 5. Log_{10} LR generated using the combined identity markers in the GlobalFiler™ PCR Amplification Kit (short tandem repeats; STRs) and ForenSeq® Kintelligence Kit (identity informative single nucleotide polymorphisms; iiSNPs) for each relationship degree for true related pairs (n = 47) and all relationship degrees for unrelated pairs (n = 152) in Pedigree Group 1. The dotted red line marks uninformative results (log_{10} LR = 0). Unrelated pairs have not been plotted for parent/offspring tests (log_{10} LR = -∞). PO: parent/offspring; FS: full sibling.

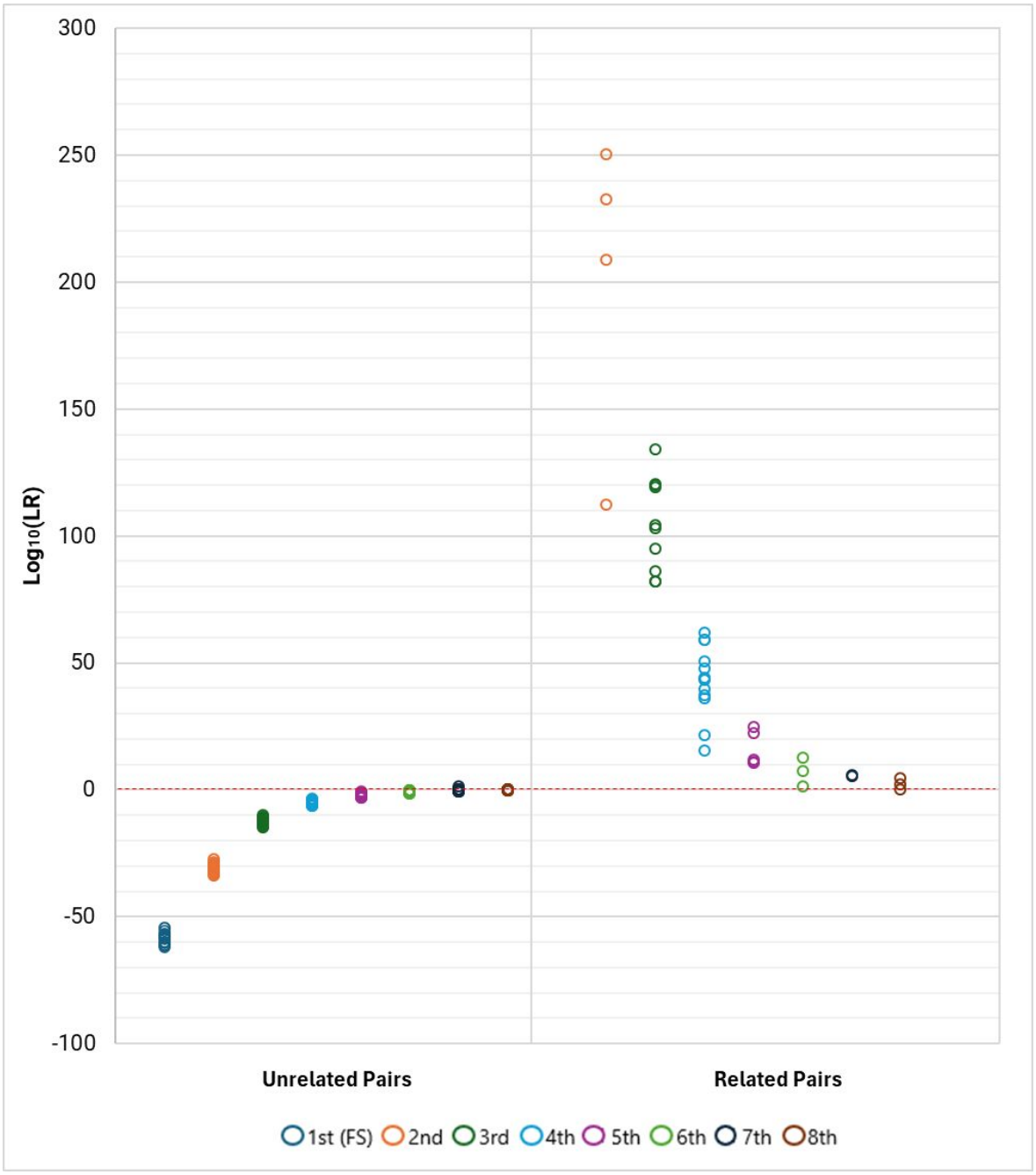

Figure 6. Log_{10} LR generated using the 10,039 autosomal single nucleotide polymorphisms (SNPs) in the ForenSeq® Kintelligence Kit for each relationship degree for true related pairs (n = 47) and all relationship degrees for unrelated pairs (n = 152) in Pedigree Group 1. The dotted red line marks uninformative results (Log_{10} LR = 0). Related parent/offspring pairs and unrelated pairs testing parent/offspring relationships have not been plotted ($\log_{10}$ LR = -∞). Related full sibling pairs have not been plotted (Log_{10} LR = ∞). FS: full sibling.

Table 7. Average log_{10} LR calculated for each relationship degree for the true related and unrelated pairs of Group 1. The range of log_{10} LRs has been given when there were fewer than five pairs tested. Parent/offspring tests were excluded ($\log_{10}$ LR = $-\infty$*).*

Relationship Degree	Related Pairs	Unrelated Pairs
1st (full siblings)	> 300	–58.61 ± 0.47
2nd	112.21 – 250.26	–31.02 ± 0.40
3rd	104.45 ± 5.49	–12.58 ± 0.29
4th	42.85 ± 3.65	–5.07 ± 0.20
5th	15.88 ± 2.77	–2.12 ± 0.16
6th	1.31 – 12.32	–0.91 ± 0.10
7th	5.16 – 5.71	–0.31 ± 0.10
8th	0.05 – 4.55	–0.12 ± 0.04

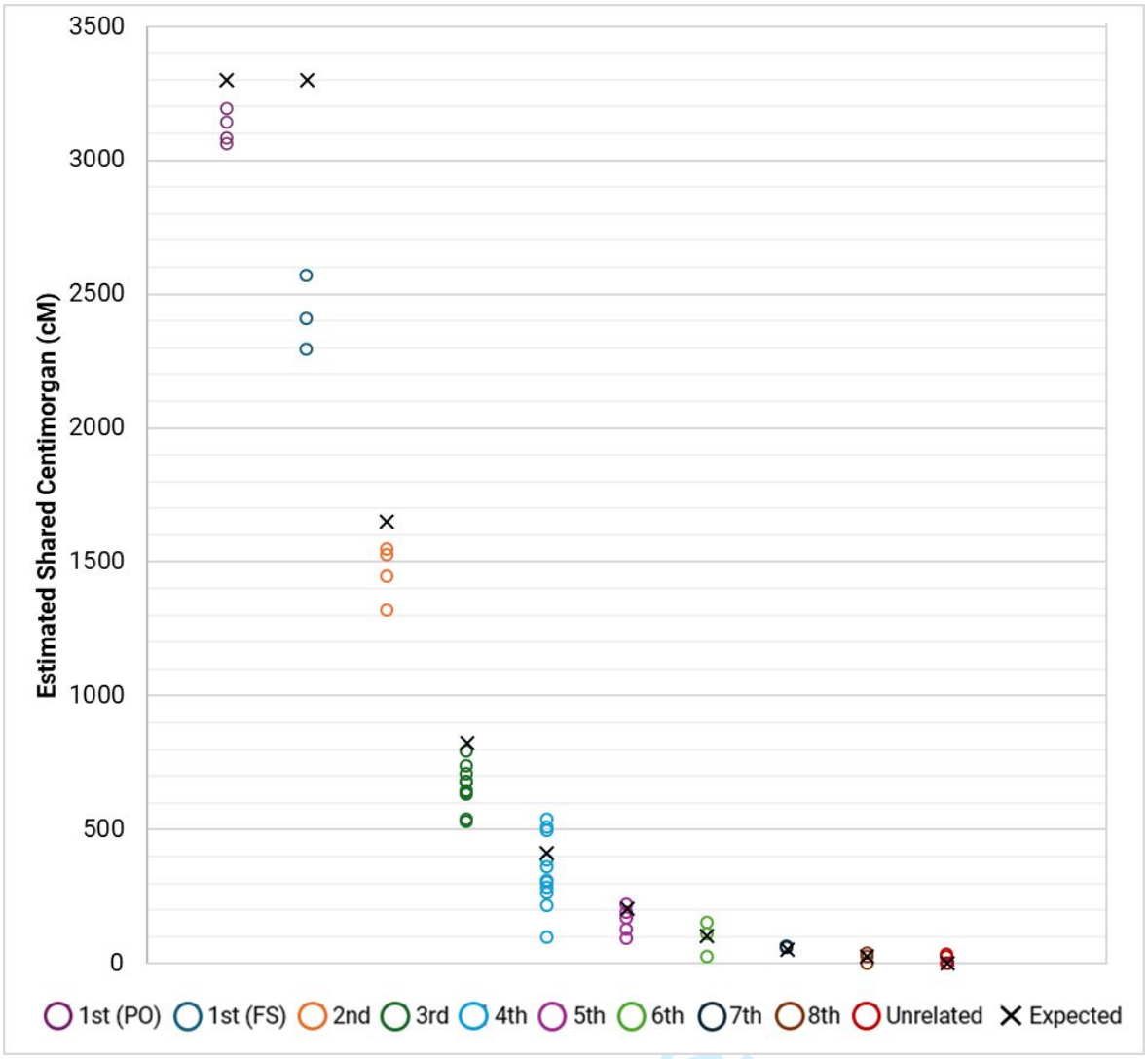

Figure 7. Estimated shared centimorgan (cM) values between the Kintelligence profiles generated by GEDmatch PRO™ for the related and unrelated pairs in Pedigree Group 1. The expected values were calculated in Table 5. PO: parent/offspring; FS: full sibling.

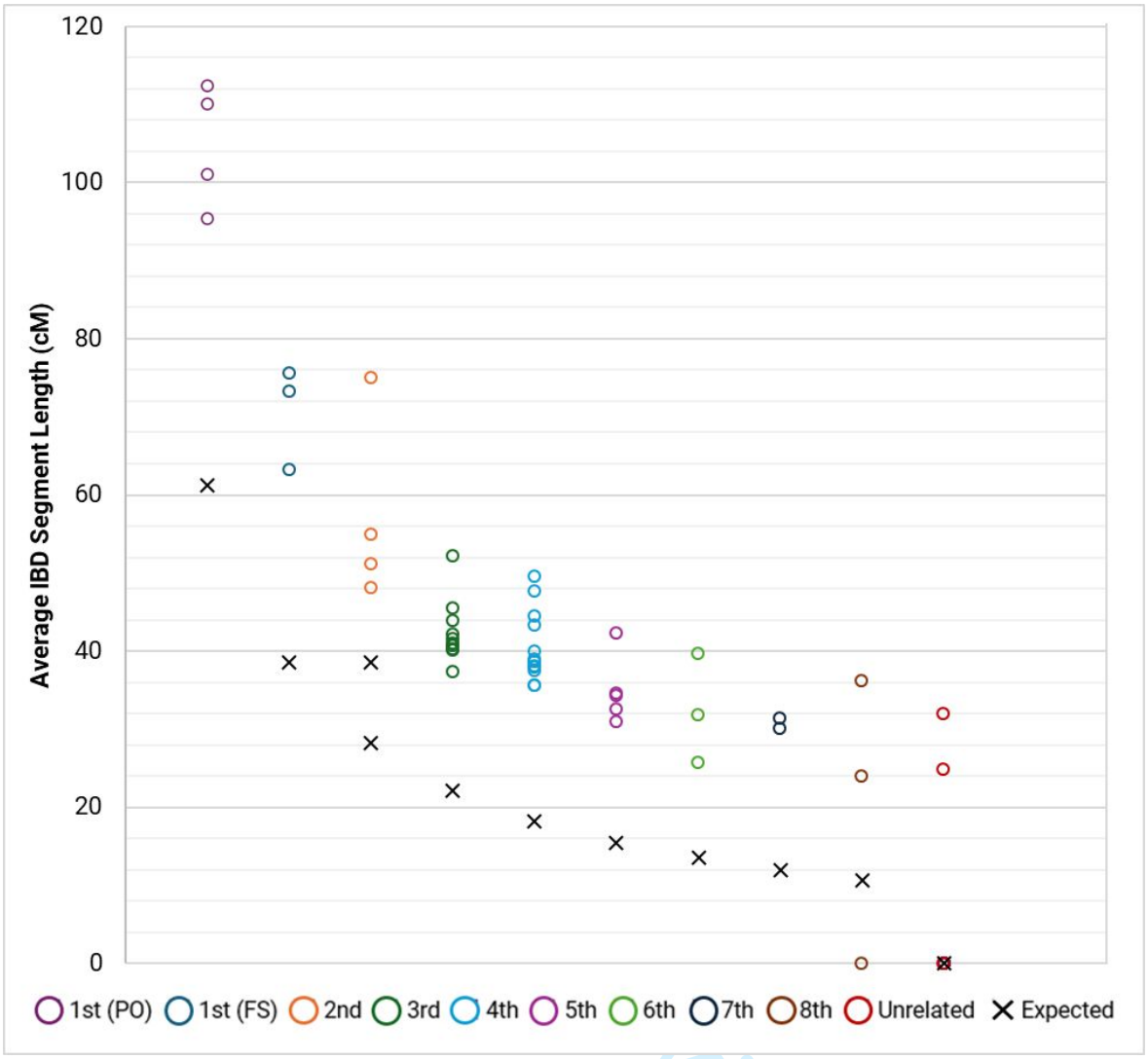

Figure 8. Average identical-by-descent (IBD) segment lengths in centimorgans (cM) in the Kintelligence profiles that were detected by GEDmatch PRO™ for the related and unrelated pairs in Pedigree Group 1. The expected values were calculated in Table 5. The two expected values for second degree relationships are different for avuncular (28.1) and grandparent/grandchild (38.6) relationships. PO: parent/offspring; FS: full sibling.

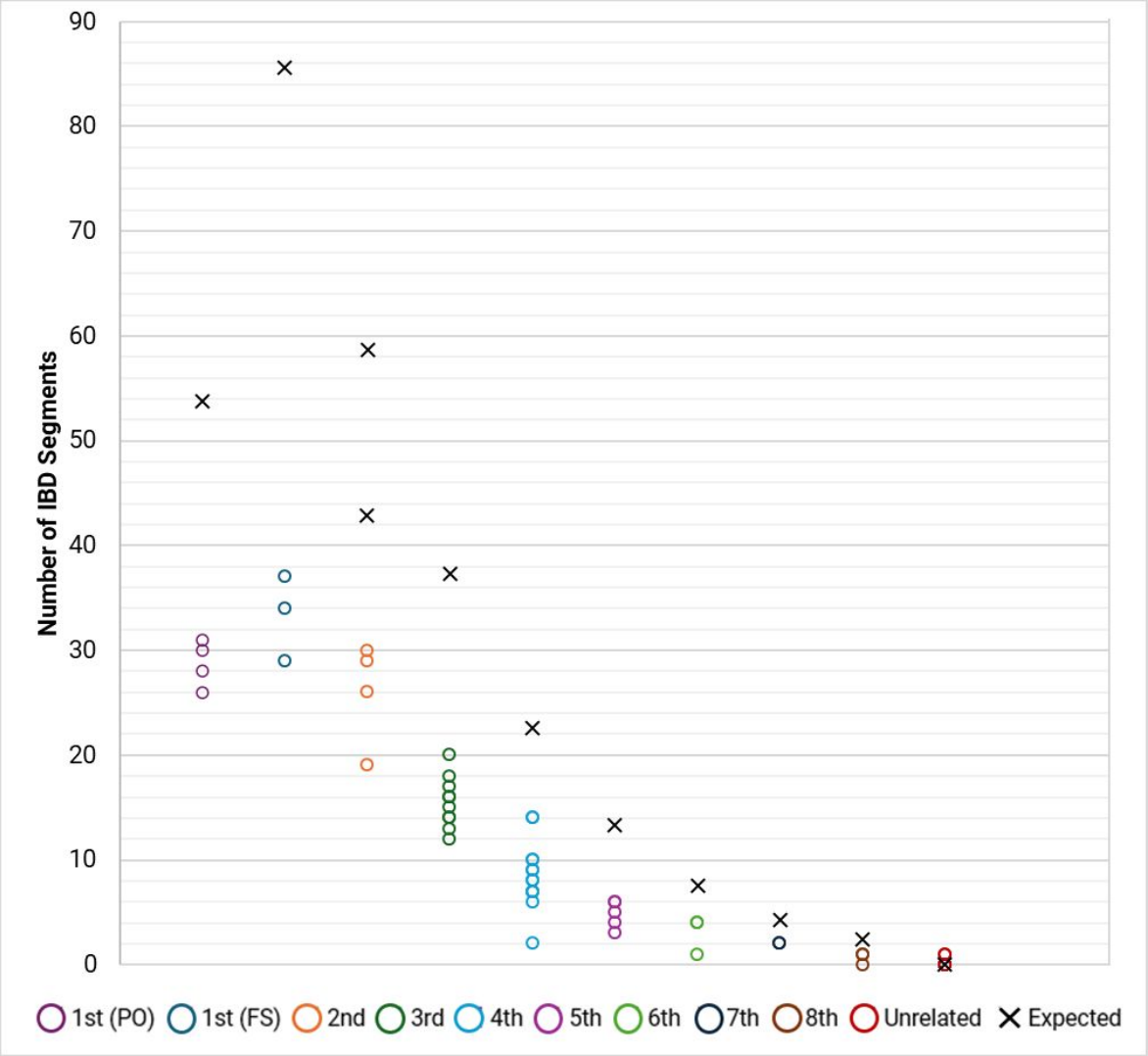

Figure 9. Number of IBD segments in the Kintelligence profiles detected by GEDmatch PRO™ for the related and unrelated pairs in Pedigree Group 1. The expected values were calculated in Table 5. The two expected values for second degree relationships are different for avuncular (58.7) and grandparent/grandchild (42.8) relationships. PO: parent/offspring; FS: full sibling.

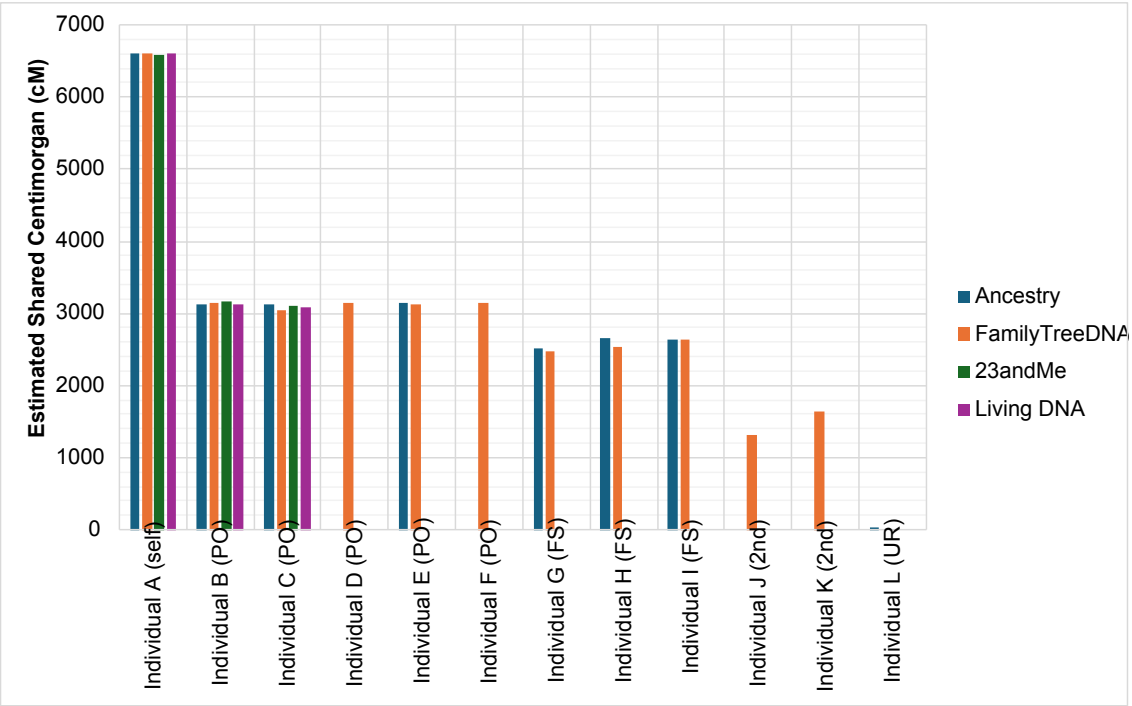

Figure 10. Estimated shared centimorgan (cM) values generated by GEDmatch PRO™ between the Kintelligence profile for Individual A and the direct-to-consumer kits from Pedigree Group 2. PO: parent/offspring; FS: full sibling; 2nd: second degree relative; UR: unrelated.

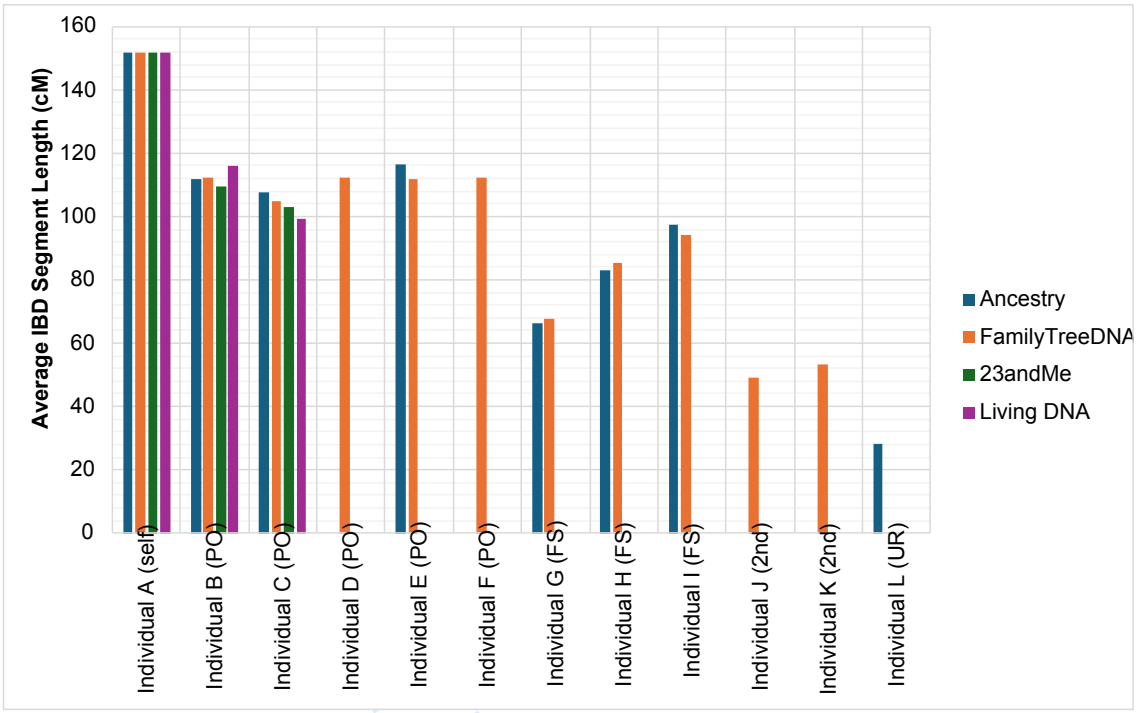

Figure 11. Average identical-by-descent (IBD) segment lengths in centimorgans (cM) detected by GEDmatch PRO™ between the Kintelligence profile for Individual A and the direct-to-consumer kits from Pedigree Group 2. PO: parent/offspring; FS: full sibling; 2nd: second degree relative; UR: unrelated.

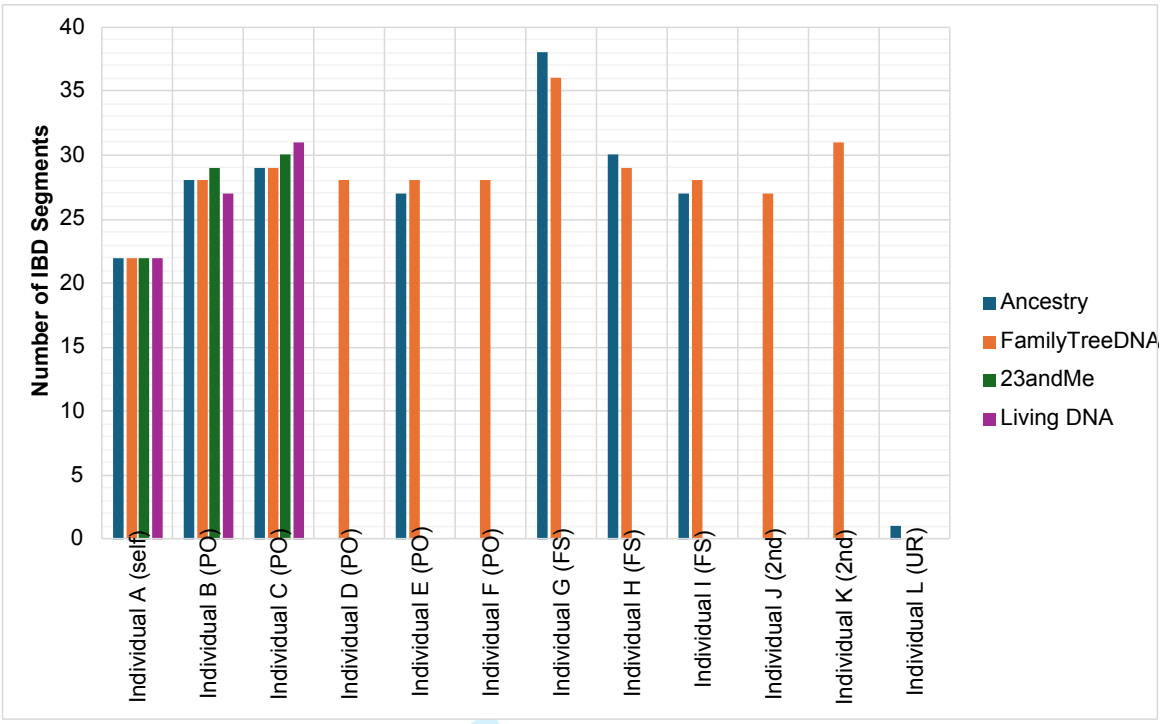

Figure 12. Number of identical-by-descent (IBD) segments detected by GEDmatch PRO™ between the Kintelligence profile for Individual A and the direct-to-consumer kits from Pedigree Group 2. PO: parent/offspring; FS: full sibling; 2nd: second degree relative; UR: unrelated.

Table 1. Direct-to-consumer kits available for Individuals A through L in Group 2.

Sample ID	AncestryDNA	FamilyTreeDNA	23andMe	Living DNA
Individual A	✓	✓	✓	✓
Individual B	✓	✓	✓	✓
Individual C	✓	✓	✓	✓
Individual D	×	✓	×	×
Individual E	✓	✓	×	×
Individual F	×	✓	×	×
Individual G	✓	✓	×	×
Individual H	✓	✓	×	×
Individual I	✓	✓	×	×
Individual J	×	✓	×	×
Individual K	×	✓	×	×
Individual L	✓	×	×	×

Table 2. Verbal scale for log_{10} LR results and the statistical support for each hypothesis.[41]

Log_{10} LR	Interpretation
≤ –5	Very strong statistical support for H_2
–4	Strong statistical support for H_2
–3	Moderately strong statistical support for H_2
–2	Moderate statistical support for H_2
–1	Limited statistical support for H_2
0	Uninformative
1	Limited statistical support for H_1
2	Moderate statistical support for H_1
3	Moderately strong statistical support for H_1
4	Strong statistical support for H_1
≥ 5	Very strong statistical support for H_1

Table 3. Population allele frequency data utilised for each marker set: short tandem repeat (STR), identity-informative single nucleotide polymorphism (iiSNP), combined identity markers (STRs and iiSNPs), and all Kintelligence SNPs.

Market Set	Panel	Markers	Population Data	Source
STRs	GlobalFiler™ PCR Amplification Kit[31]	21	Australian Caucasian	Taylor *et al.* (2017)[43]
iiSNPs	ForenSeq® Kintelligence Kit[34]	94	Australian with European Ancestry	Watson *et al.* (2024)[44]

Combined Identity Markers	GlobalFiler™ PCR Amplification Kit and ForenSeq® Kintelligence Kit[31, 34]	114[a]	Australian Caucasian (STR) Australian with European Ancestry (iiSNP)	Taylor *et al.* (2017) and Watson *et al.* (2024)[43, 44]
Kintelligence SNPs	ForenSeq® Kintelligence Kit[34]	10,039[b]	European	1000 Genomes[45]

[a] SE33 was excluded as linkage equilibrium tests could not be performed between SE33 and iiSNPs.[44]

[b] X and Y SNPs were excluded from the Kintelligence Kit for likelihood ratio calculations.

Table 4. Shared centimorgan (cM) thresholds applied in the GEDmatch PRO™ One-to-Many Kinship tool match lists based on the number of overlapping SNPs. For all threshold levels, the longest identical-by-descent (IBD) segment needs to be at least 30 cM.[17]

Overlapping SNPs	Shared cM Threshold	
	High Confidence Match List	Expanded Match List
9000	170	120
8000	190	140
6000	200	160

Table 5. Expected values for total shared centimorgan (cM), average identical by descent (IBD) segment length and number of IBD segments calculated for different relationships. Values were calculated with the equations in Supplementary Information S1.

Kinship Degree	Relationship	Total Shared cM	Average IBD Segment Length (cM)	Number of IBD Segments
1	Parent / offspring	3300	61.3	53.8
1	Full sibling	3300	38.6	85.6
2	Avuncular	1650	28.1	58.7
2	Grandparent / grandchild	1650	38.6	42.8
3	First cousin	825	22.1	37.3
4	First cousin once removed	412.5	18.2	22.6
5	Second cousin	206.3	15.5	13.3
6	Second cousin once removed	103.1	13.5	7.6
7	Second cousin twice removed	51.6	11.9	4.3
8	Third cousin once removed	25.8	10.7	2.4
N/A	Unrelated	0	0	0

Table 6. Call rates for the DNA profiles generated for each sample in Pedigree Group 1; calculated for short tandem repeats (STRs; 21), identity-informative SNPs (iiSNPs; 94), combined identity markers (STRs and iiSNPs; 114) and autosomal Kintelligence SNPs (10,039).

Sample ID	STRs (%)	iiSNPs (%)	Combined Identity Markers (114)	Kintelligence (%)
Individual 1	100.00	100.00	100.00	99.86
Individual 2	100.00	100.00	100.00	99.82
Individual 3	100.00	96.81	98.25	96.96
Individual 4	100.00	95.74	96.49	97.85
Individual 5	85.71	97.87	97.80	96.18
Individual 6	100.00	100.00	100.00	99.76

Individual 7	100.00	98.94	99.12	99.76
Individual 8	100.00	100.00	100.00	99.95
Individual 9	100.00	98.94	99.12	99.74
Individual 10	95.24	92.55	93.86	92.02
Individual 11	100.00	100.00	100.00	99.79
Individual 12	100.00	98.94	99.12	99.41

Table 7. Average log_{10} LR calculated for each relationship degree for the true related and unrelated pairs of Group 1. The range of log_{10} LRs has been given when there were fewer than five pairs tested. Parent/offspring tests were excluded ($\log_{10}$ LR = $-\infty$*).*

Relationship Degree	Related Pairs	Unrelated Pairs
1st (full siblings)	> 300	–58.61 ± 0.47
2nd	112.21 – 250.26	–31.02 ± 0.40
3rd	104.45 ± 5.49	–12.58 ± 0.29
4th	42.85 ± 3.65	–5.07 ± 0.20
5th	15.88 ± 2.77	–2.12 ± 0.16
6th	1.31 – 12.32	–0.91 ± 0.10
7th	5.16 – 5.71	–0.31 ± 0.10
8th	0.05 – 4.55	–0.12 ± 0.04

Supplementary Information S1: Calculations for expected shared centimorgan (cM), average identical by descent (IBD) segment length and number of IBD segments. Equations are modified from those published by Speed & Balding (2015) for total length of DNA in cM.[1]

The average IBD segment length that two individuals share is:

$$\text{Average IBD Segment Length} = \frac{3300}{22 + (40.7 + 22.9) \times \frac{G_1 + G_2}{2}} \quad (1)$$

where the numerator is the total length of DNA in cM; A is the number of most recent common ancestors shared between the individuals; G_1 and G_2 are the number of generations between each individual and the most recent common ancestor(s); 22 is the number of autosomal chromosomes; and $(40.7 + 22.9)/2$ is the sex-averaged number of recombinations per meiosis.

The total cM that two individuals share is:

$$\text{Expected Shared cM} = A \times \frac{1}{2^{G_1+G_2-1}} \times 3300 \quad (2)$$

The number of IBD segments that two individuals share is:

$$\text{Number of IBD Segments} = A \times \frac{22 + (40.7 + 22.9) \times \frac{G_1 + G_2}{2}}{2^{G_1+G_2-1}} \quad (3)$$

Reference

1. Speed, D., Balding, D. Relatedness in the post-genomic era: is it still useful? Nat Rev Genet 2015;16,:33–44 (2015).; doi: 10.1038/nrg3821.

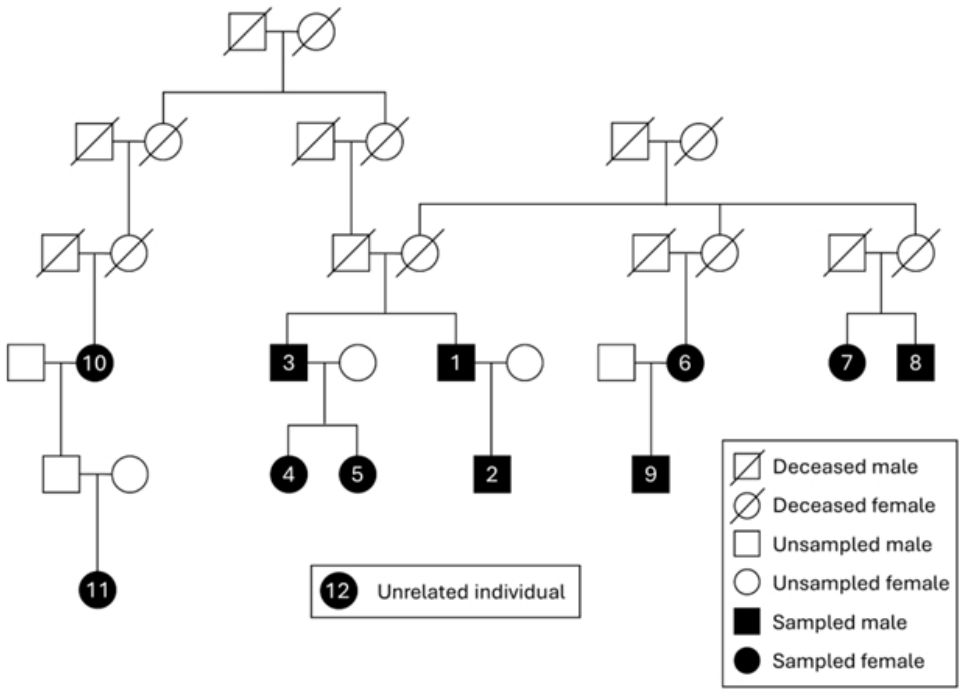

Figure 1. Pedigree Group 1, with numbers corresponding to Individuals 1 through 12. Individuals included in the study are filled in (black), with living relatives linking the family members not filled (white). Deceased relatives are crossed out. Individual 12 is unrelated to all other individuals in the group.

169x121mm (96 x 96 DPI)

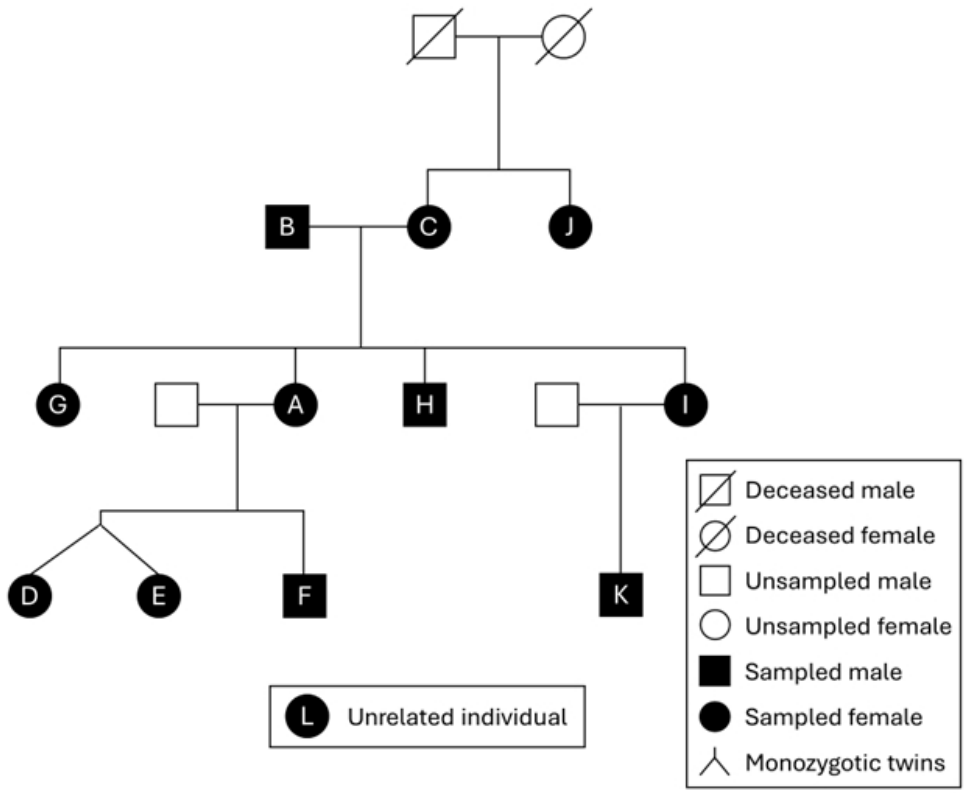

Figure 2. Pedigree Group 2, with letters corresponding to Individuals A through L. Individuals included in the study are filled in (black), with living relatives linking the family members not filled (white). Deceased relatives are crossed out. Individual L is unrelated to all other individuals in the group.

169x139mm (96 x 96 DPI)

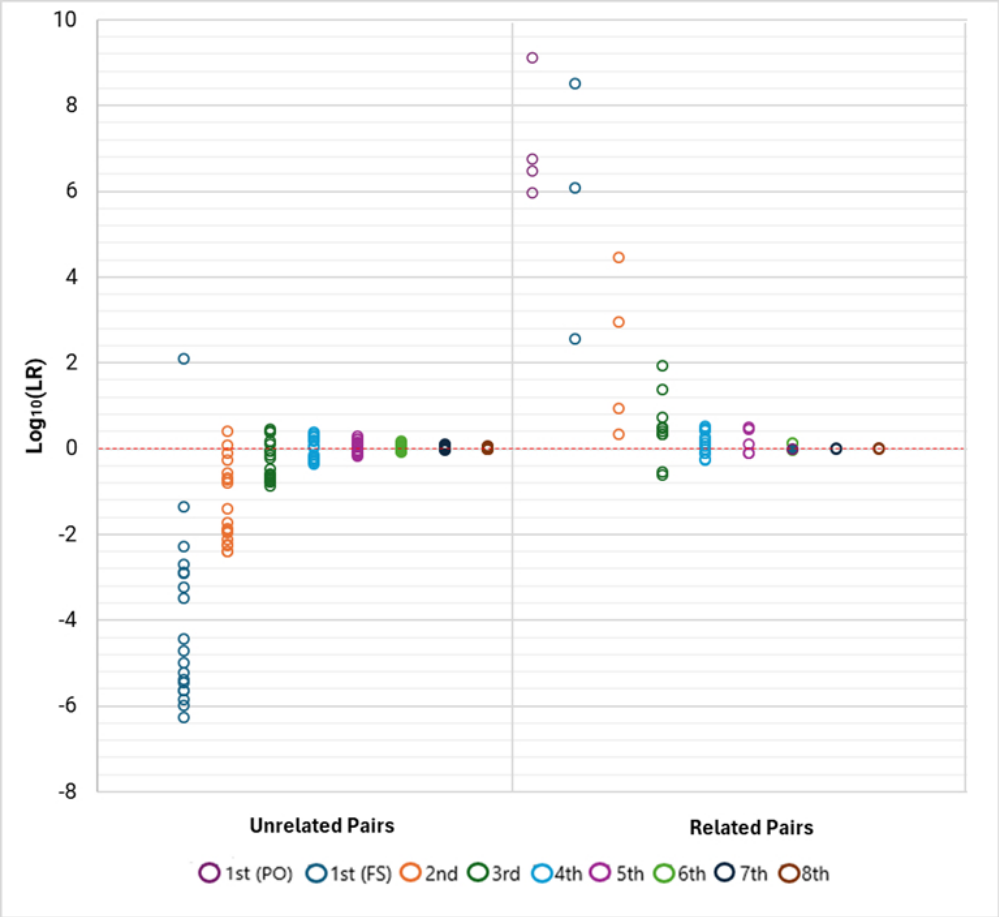

Figure 3. Log10 LR generated using the short tandem repeats (STRs) in the GlobalFiler™ PCR Amplification Kit for each relationship degree for true related pairs (n = 47) and all relationship degrees for unrelated pairs (n = 152) in Pedigree Group 1. The dotted red line marks uninformative results (log10 LR = 0). Unrelated pairs have not been plotted for parent/offspring tests (log10 LR = -∞). PO: parent/offspring; FS: full sibling.

130x119mm (150 x 150 DPI)

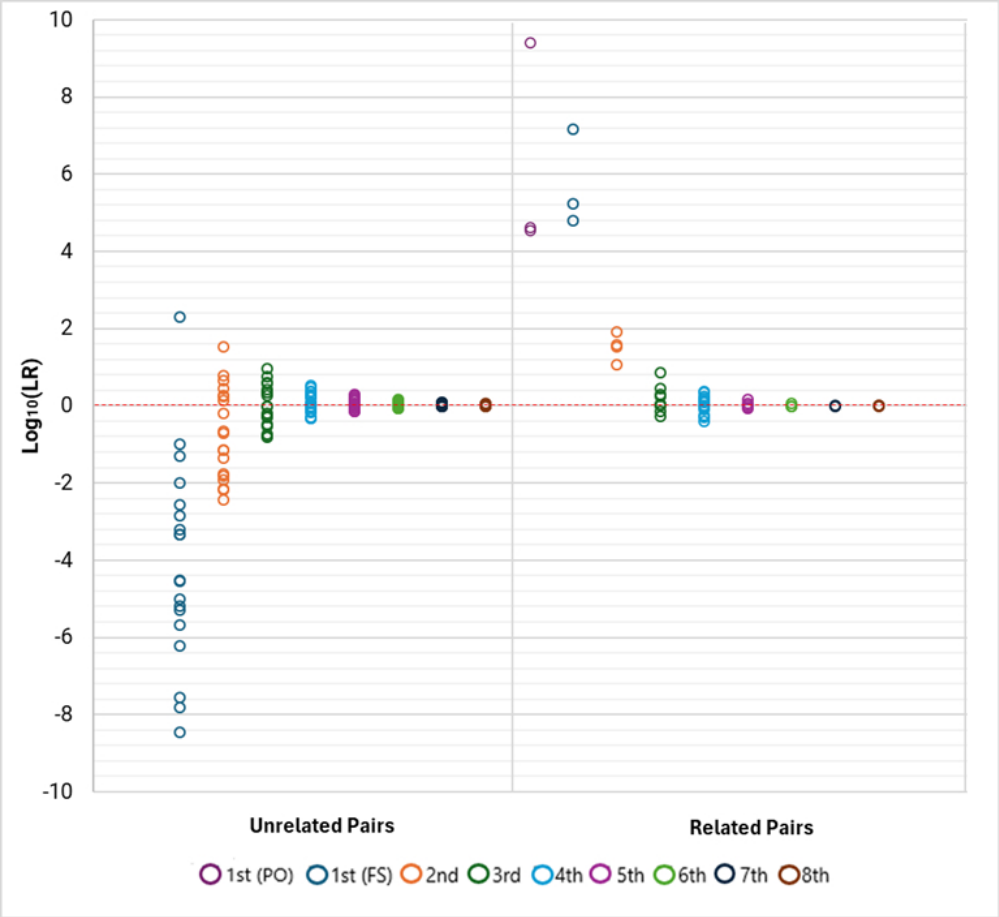

Figure 4. Log10 LR generated using the identity-informative single nucleotide polymorphisms (iiSNPs) in the ForenSeq® Kintelligence Kit for each relationship degree for true related pairs (n = 47) and all relationship degrees for unrelated pairs (n = 152) in Pedigree Group 1. The dotted red line marks uninformative results (log10 LR = 0). Unrelated pairs have not been plotted for parent/offspring tests (log10 LR = -∞). PO: parent/offspring; FS: full sibling.

130x119mm (150 x 150 DPI)

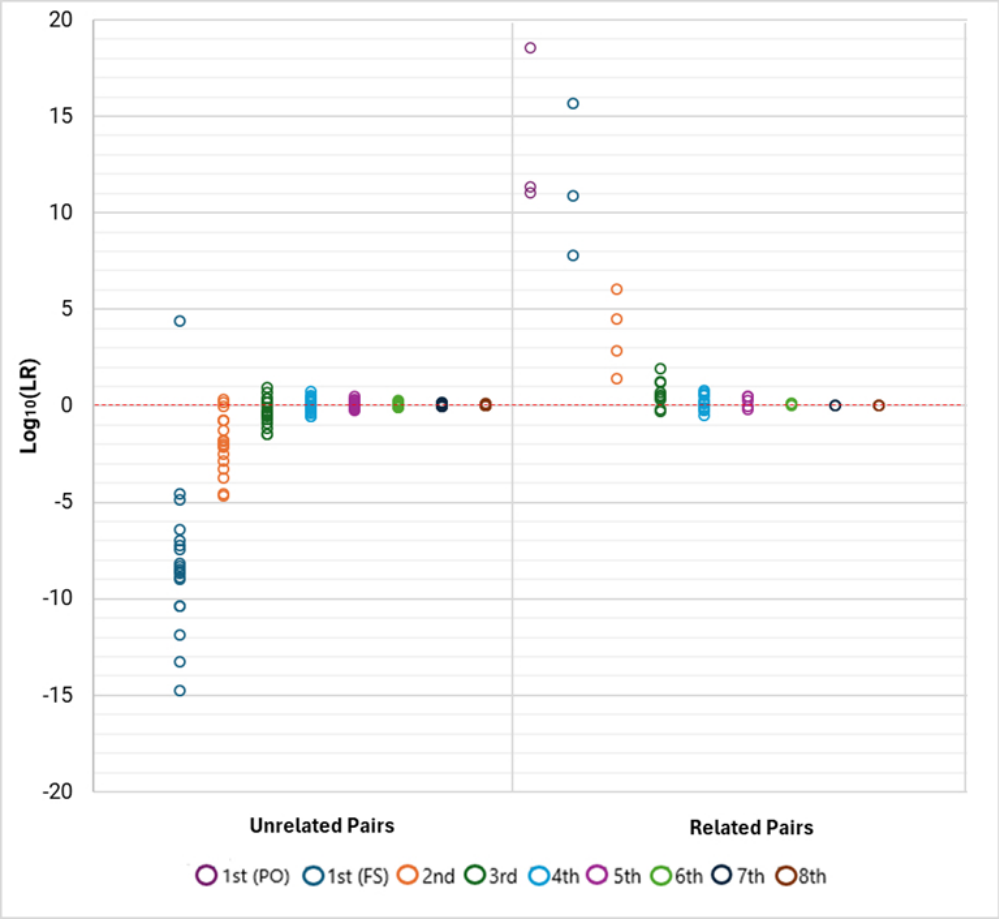

Figure 5. Log10 LR generated using the combined identity markers in the GlobalFiler™ PCR Amplification Kit (short tandem repeats; STRs) and ForenSeq® Kintelligence Kit (identity informative single nucleotide polymorphisms; iiSNPs) for each relationship degree for true related pairs (n = 47) and all relationship degrees for unrelated pairs (n = 152) in Pedigree Group 1. The dotted red line marks uninformative results (log10 LR = 0). Unrelated pairs have not been plotted for parent/offspring tests (log10 LR = -∞). PO: parent/offspring; FS: full sibling.

130x119mm (150 x 150 DPI)

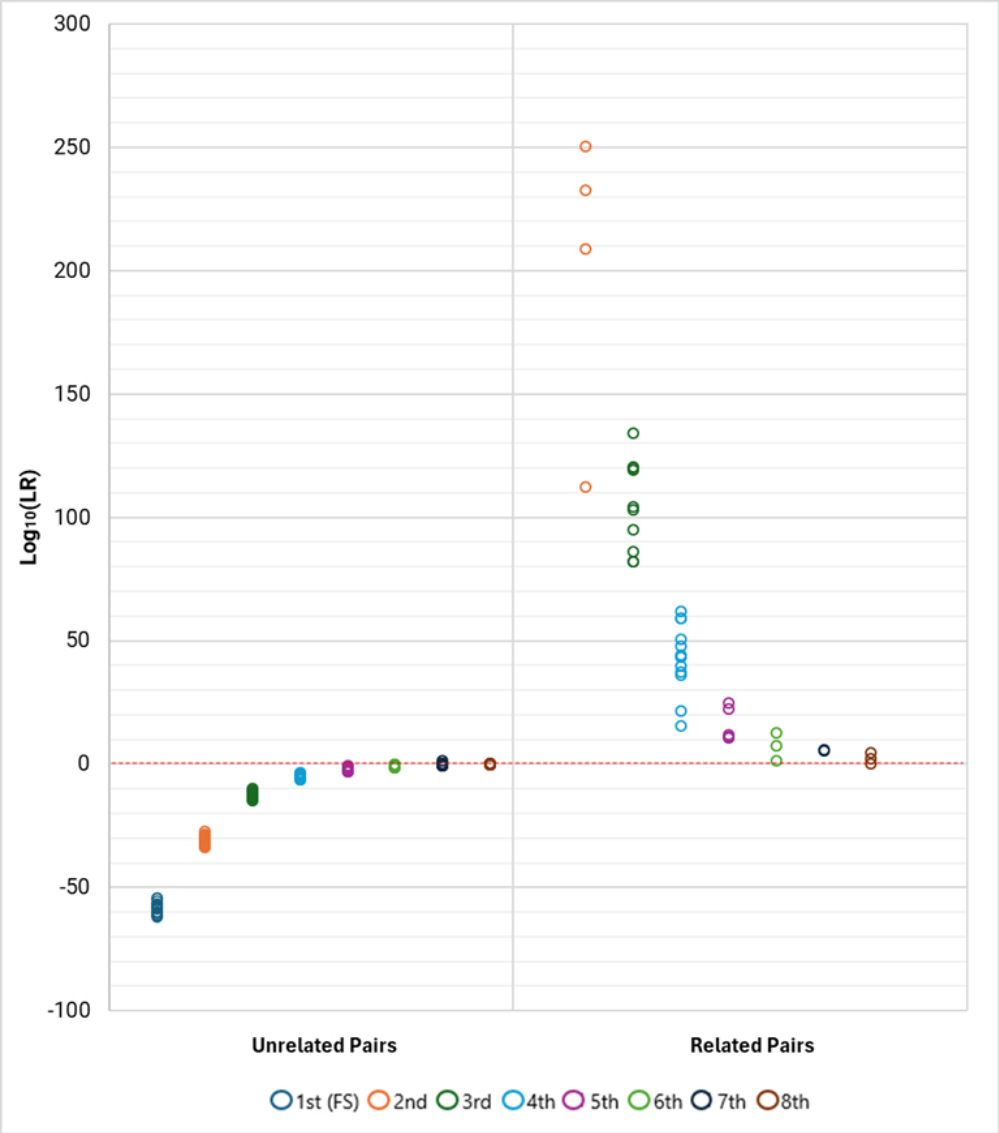

Figure 6. Log10 LR generated using the 10,039 autosomal single nucleotide polymorphisms (SNPs) in the ForenSeq® Kintelligence Kit for each relationship degree for true related pairs (n = 47) and all relationship degrees for unrelated pairs (n = 152) in Pedigree Group 1. The dotted red line marks uninformative results (Log10 LR = 0). Related parent/offspring pairs and unrelated pairs testing parent/offspring relationships have not been plotted (log10 LR = -∞). Related full sibling pairs have not been plotted (Log10 LR = ∞). FS: full sibling.

130x147mm (150 x 150 DPI)

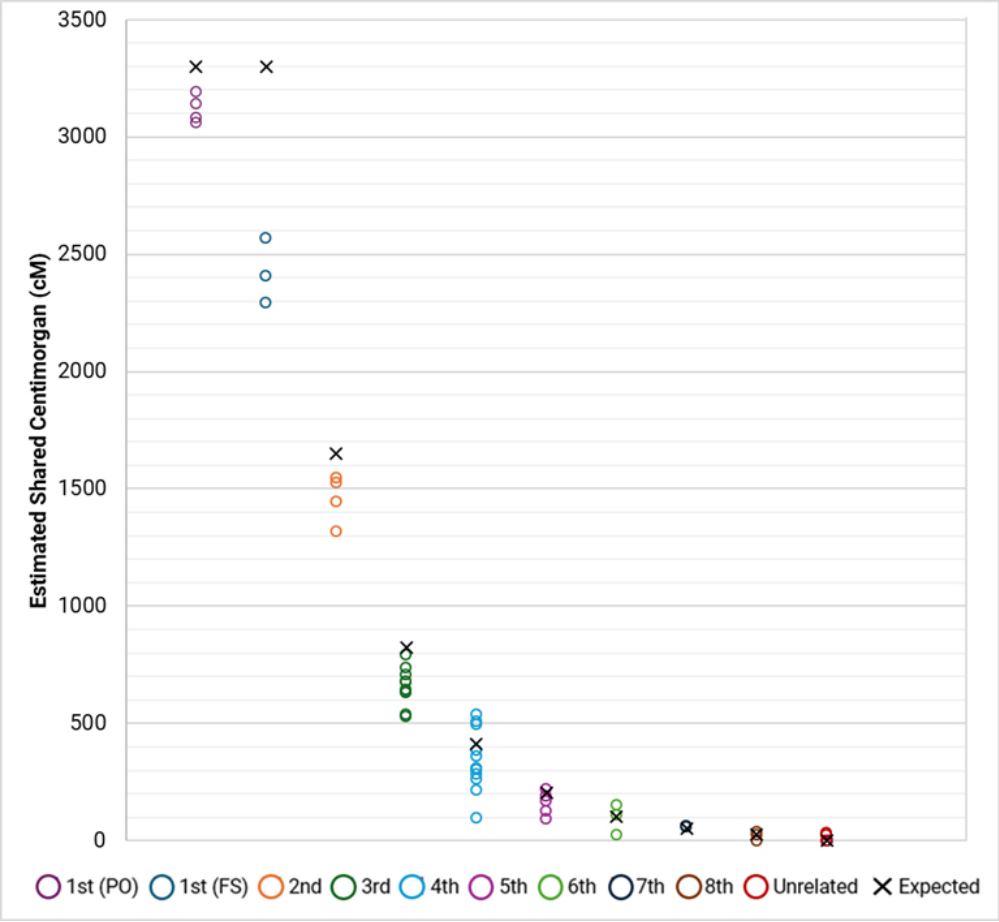

Figure 7. Estimated shared centimorgan (cM) values between the Kintelligence profiles generated by GEDmatch PRO™ for the related and unrelated pairs in Pedigree Group 1. The expected values were calculated in Table 5. PO: parent/offspring; FS: full sibling.

130x119mm (150 x 150 DPI)

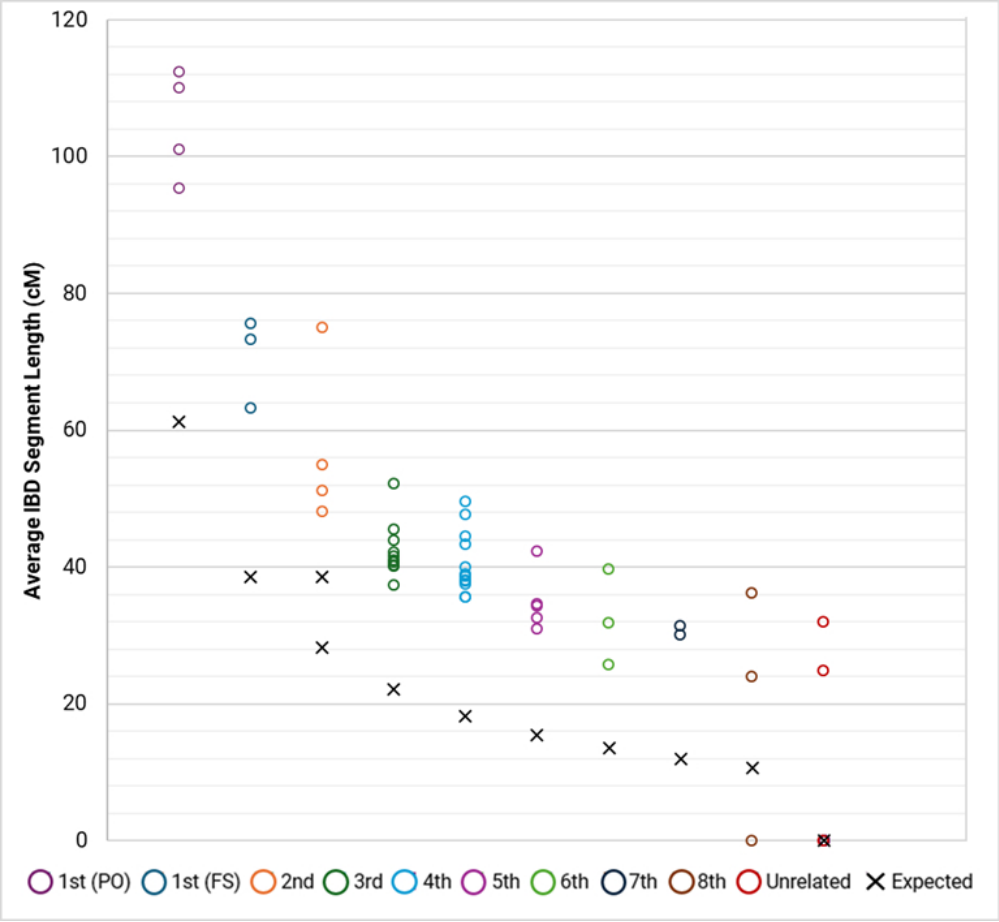

Figure 8. Average identical-by-descent (IBD) segment lengths in centimorgans (cM) in the Kintelligence profiles that were detected by GEDmatch PRO™ for the related and unrelated pairs in Pedigree Group 1. The expected values were calculated in Table 5. The two expected values for second degree relationships are different for avuncular (28.1) and grandparent/grandchild (38.6) relationships. PO: parent/offspring; FS: full sibling.

130x119mm (150 x 150 DPI)

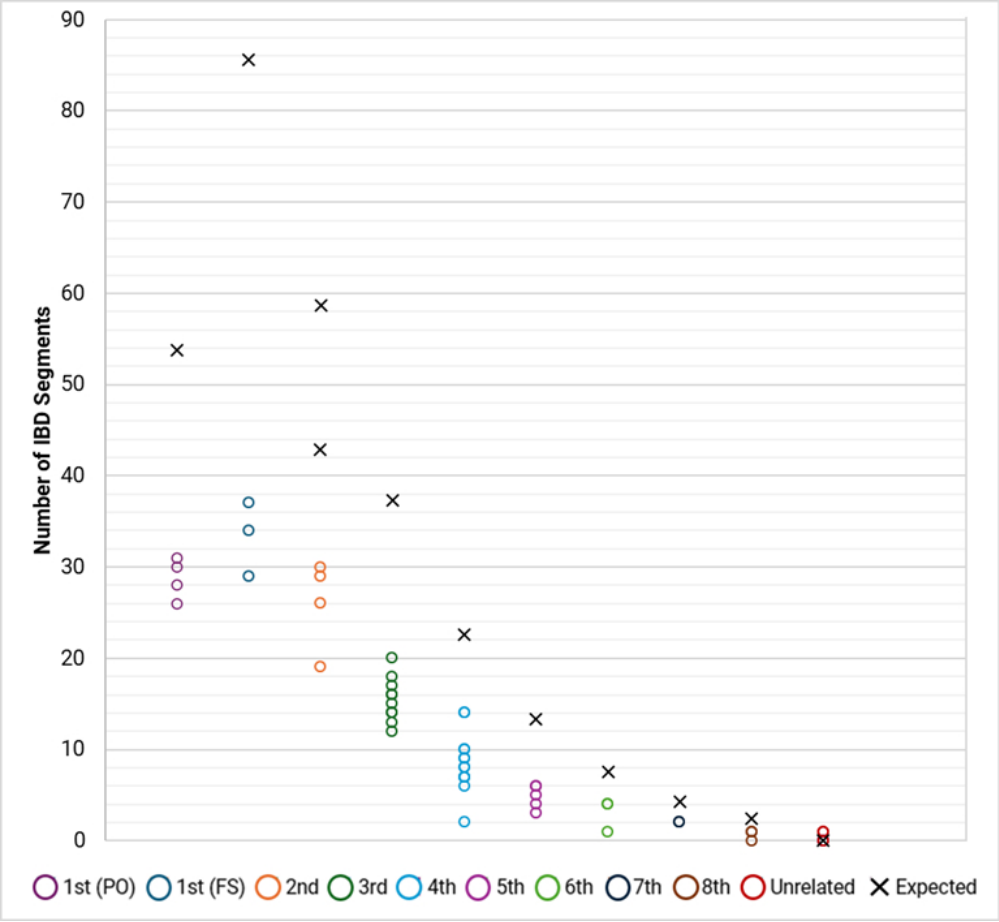

Figure 9. Number of IBD segments in the Kintelligence profiles detected by GEDmatch PRO™ for the related and unrelated pairs in Pedigree Group 1. The expected values were calculated in Table 5. The two expected values for second degree relationships are different for avuncular (58.7) and grandparent/grandchild (42.8) relationships. PO: parent/offspring; FS: full sibling.

130x119mm (150 x 150 DPI)

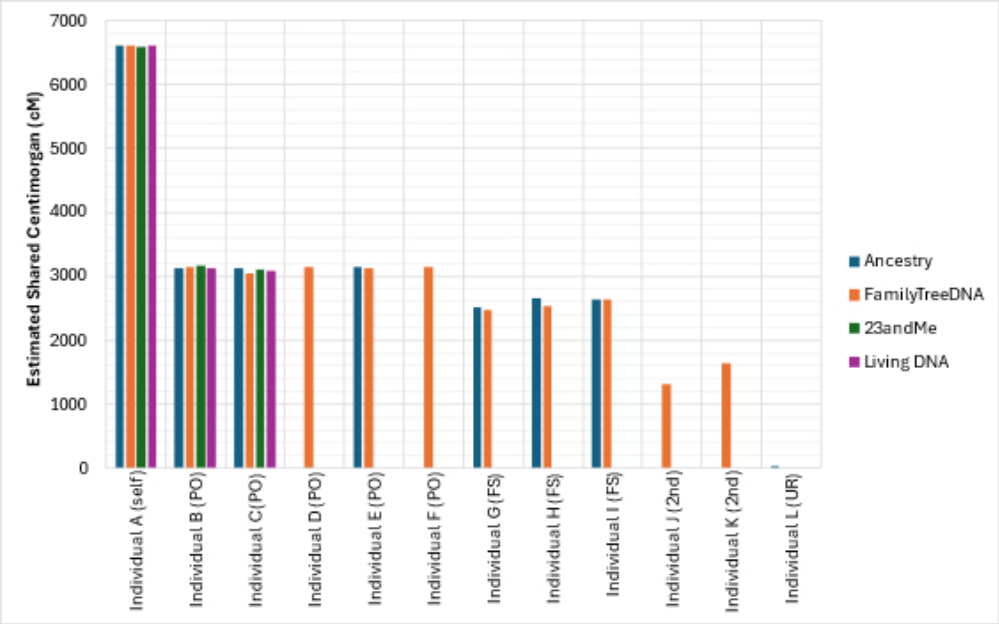

Figure 10. Estimated shared centimorgan (cM) values generated by GEDmatch PRO™ between the Kintelligence profile for Individual A and the direct-to-consumer kits from Pedigree Group 2. PO: parent/offspring; FS: full sibling; 2nd: second degree relative; UR: unrelated.

149x93mm (96 x 96 DPI)

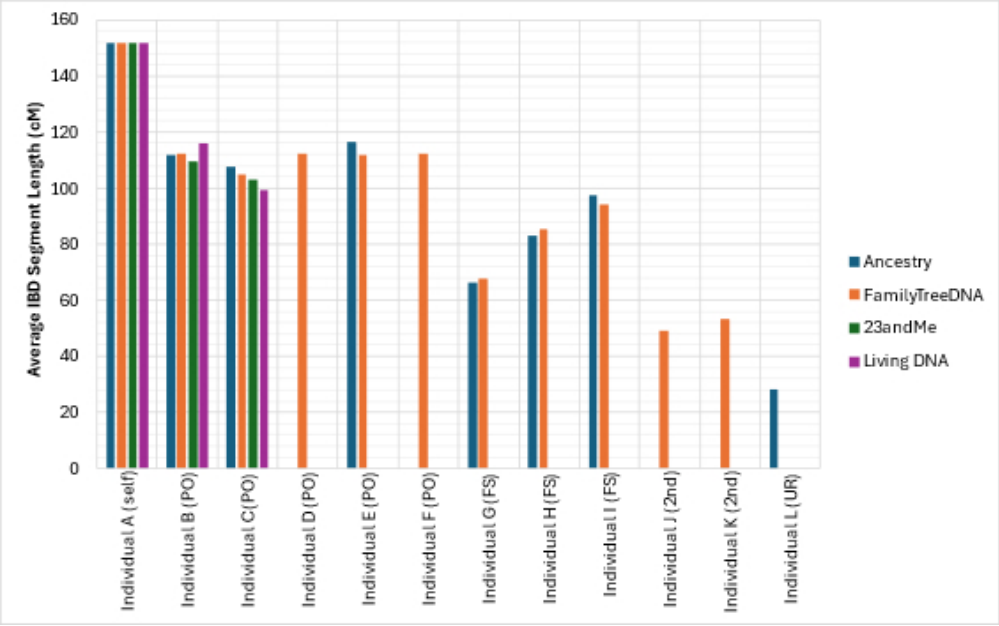

Figure 11. Average identical-by-descent (IBD) segment lengths in centimorgans (cM) detected by GEDmatch PRO™ between the Kintelligence profile for Individual A and the direct-to-consumer kits from Pedigree Group 2. PO: parent/offspring; FS: full sibling; 2nd: second degree relative; UR: unrelated.

149x93mm (96 x 96 DPI)

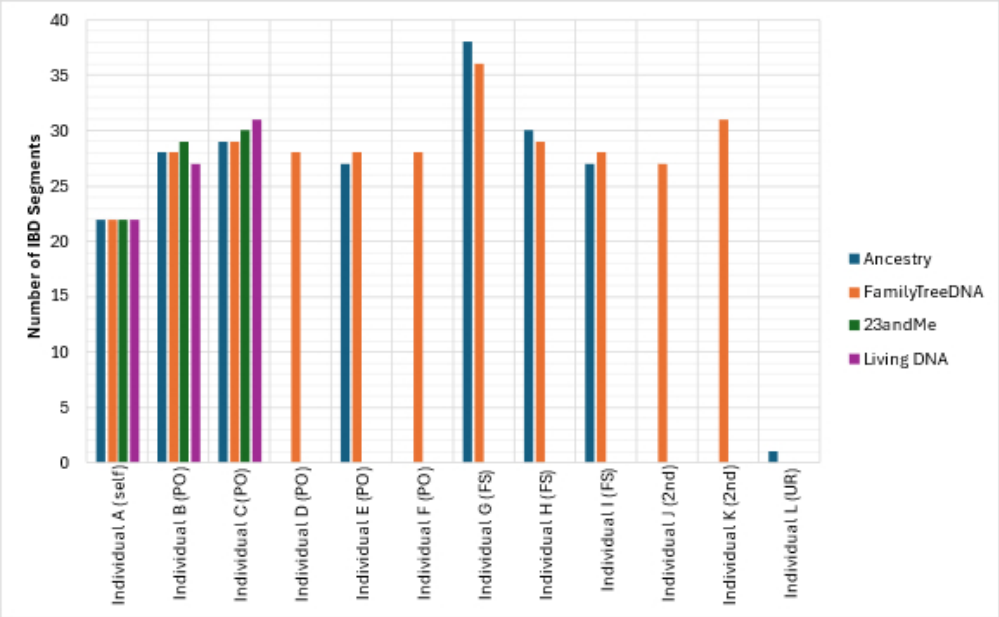

Figure 12. Number of identical-by-descent (IBD) segments detected by GEDmatch PRO™ between the Kintelligence profile for Individual A and the direct-to-consumer kits from Pedigree Group 2. PO: parent/offspring; FS: full sibling; 2nd: second degree relative; UR: unrelated.

152x94mm (96 x 96 DPI)

Figure 1. Pedigree Group 1, with numbers corresponding to Individuals 1 through 12. Individuals included in the study are filled in (black), with living relatives linking the family members not filled (white). Deceased relatives are crossed out. Individual 12 is unrelated to all other individuals in the group.

Figure 2. Pedigree Group 2, with letters corresponding to Individuals A through L. Individuals included in the study are filled in (black), with living relatives linking the family members not filled (white). Deceased relatives are crossed out. Individual L is unrelated to all other individuals in the group.

Figure 3. Log_{10} LR generated using the short tandem repeats (STRs) in the GlobalFiler™ PCR Amplification Kit for each relationship degree for true related pairs (n = 47) and all relationship degrees for unrelated pairs (n = 152) in Pedigree Group 1. The dotted red line marks uninformative results (log_{10} LR = 0). Unrelated pairs have not been plotted for parent/offspring tests (log_{10} LR = $-\infty$). PO: parent/offspring; FS: full sibling.

Figure 4. Log_{10} LR generated using the identity-informative single nucleotide polymorphisms (iiSNPs) in the ForenSeq® Kintelligence Kit for each relationship degree for true related pairs (n = 47) and all relationship degrees for unrelated pairs (n = 152) in Pedigree Group 1. The dotted red line marks uninformative results (log_{10} LR = 0). Unrelated pairs have not been plotted for parent/offspring tests (log_{10} LR = $-\infty$). PO: parent/offspring; FS: full sibling.

Figure 5. Log_{10} LR generated using the combined identity markers in the GlobalFiler™ PCR Amplification Kit (short tandem repeats; STRs) and ForenSeq® Kintelligence Kit (identity informative single nucleotide polymorphisms; iiSNPs) for each relationship degree for true related pairs (n = 47)

and all relationship degrees for unrelated pairs (n = 152) in Pedigree Group 1. The dotted red line marks uninformative results (log_{10} LR = 0). Unrelated pairs have not been plotted for parent/offspring tests (log_{10} LR = -∞). PO: parent/offspring; FS: full sibling.

Figure 6. Log_{10} LR generated using the 10,039 autosomal single nucleotide polymorphisms (SNPs) in the ForenSeq® Kintelligence Kit for each relationship degree for true related pairs (n = 47) and all relationship degrees for unrelated pairs (n = 152) in Pedigree Group 1. The dotted red line marks uninformative results (Log_{10} LR = 0). Related parent/offspring pairs and unrelated pairs testing parent/offspring relationships have not been plotted (log_{10} LR = -∞). Related full sibling pairs have not been plotted (Log_{10} LR = ∞). FS: full sibling.

Figure 7. Estimated shared centimorgan (cM) values between the Kintelligence profiles generated by GEDmatch PRO™ for the related and unrelated pairs in Pedigree Group 1. The expected values were calculated in Table 5. PO: parent/offspring; FS: full sibling.

Figure 8. Average identical-by-descent (IBD) segment lengths in centimorgans (cM) in the Kintelligence profiles that were detected by GEDmatch PRO™ for the related and unrelated pairs in Pedigree Group 1. The expected values were calculated in Table 5. The two expected values for second degree relationships are different for avuncular (28.1) and grandparent/grandchild (38.6) relationships. PO: parent/offspring; FS: full sibling.

Figure 9. Number of IBD segments in the Kintelligence profiles detected by GEDmatch PRO™ for the related and unrelated pairs in Pedigree Group 1. The expected values were calculated in Table 5. The

two expected values for second degree relationships are different for avuncular (58.7) and grandparent/grandchild (42.8) relationships. PO: parent/offspring; FS: full sibling.

Figure 10. Estimated shared centimorgan (cM) values generated by GEDmatch PRO™ between the Kintelligence profile for Individual A and the direct-to-consumer kits from Pedigree Group 2. PO: parent/offspring; FS: full sibling; 2nd: second degree relative; UR: unrelated.

Figure 11. Average identical-by-descent (IBD) segment lengths in centimorgans (cM) detected by GEDmatch PRO™ between the Kintelligence profile for Individual A and the direct-to-consumer kits from Pedigree Group 2. PO: parent/offspring; FS: full sibling; 2nd: second degree relative; UR: unrelated.

Figure 12. Number of identical-by-descent (IBD) segments detected by GEDmatch PRO™ between the Kintelligence profile for Individual A and the direct-to-consumer kits from Pedigree Group 2. PO: parent/offspring; FS: full sibling; 2nd: second degree relative; UR: unrelated.