Impact of SNP microarray analysis of compromised DNA on kinship classification success in the context of investigative genetic genealogy

Short title: SNP microarrays on compromised DNA for genetic genealogy

Jard H. de Vries[1,§], Daniel Kling[2,§], Athina Vidaki[3], Pascal Arp[1], Vivian Kalamara[3], Michael M.P.J. Verbiest[1], Danuta Piniewska-Róg[4,5], Thomas J. Parsons[6], André G. Uitterlinden[1,7], Manfred Kayser[3,*]

[1] Erasmus MC, University Medical Center Rotterdam, Department of Internal Medicine, Dr. Molewaterplein 40, 3015 GD, Rotterdam, the Netherlands

[2] Department of Forensic Genetics and Toxicology, National Board of Forensic Medicine, Artillerigatan 12, 587 58 Linköping, Sweden

[3] Erasmus MC, University Medical Center Rotterdam, Department of Genetic Identification, Dr. Molewaterplein 40, 3015 GD, Rotterdam, the Netherlands

[4] Malopolska Centre of Biotechnology, Jagiellonian University, 30-387 Krakow, Poland

[5] Department of Forensic Medicine, Jagiellonian University Medical College, 31-531 Krakow, Poland

1

[6] International Commission on Missing Persons, Koninginnegracht 12a, 2514 AA, Den Haag, the Netherlands

[7] Erasmus MC, University Medical Center Rotterdam, Department of Epidemiology, Dr. Molewaterplein 40, 3015 GD, Rotterdam, the Netherlands

[§]these authors contributed equally to this work.

*To whom correspondence may be addressed. Prof. Dr. Manfred Kayser, Email: m.kayser@erasmusmc.nl

ORCID IDs

J.H.V. https://orcid.org/0000-0001-8534-8994;

D.K. https://orcid.org/0000-0002-4157-7194;

A.V. https://orcid.org/0000-0002-5470-245X;

V.K. https://orcid.org/0000-0002-9313-4109;

D.P.-R. https://orcid.org/0000-0001-8307-9145;

A.G.U. https://orcid.org/0000-0002-7276-3387;

M.K. https://orcid.org/0000-0002-4958-847X;

Author Contributions

J.H.V., D.K., A.V., A.G.U., M.K. designed study with contributions from T.J.P.;

J.H.V., D.K., A.V., P.A., V.K., M.M.P.J.V. generated data;

J.H.V., D.K., A.V. analyzed data;

A.V. prepared main display items with contributions from J.H.V. and D.K.;

A.G.U., D.P.-R., M.K. contributed resources;

J.H.V., D.K., A.V., M.K. wrote the manuscript;

P.A., V.K., M.M.P.J.V., D.P.-R., T.J.P., A.G.U. edited versions of the manuscript;

All authors approved the final manuscript.

Competing Interest Statement:

There were no competing interests.

1 **Abstract**

2 Single nucleotide polymorphism (SNP) data generated with microarray technologies have been

3 used to solve murder cases via investigative leads obtained from identifying relatives of the

4 unknown perpetrator included in accessible genomic databases, referred to as investigative

5 genetic genealogy (IGG). However, SNP microarrays were developed for relatively high input

6 DNA quantity and quality, while SNP microarray data from compromised DNA typically

7 obtainable from crime scene stains are largely missing. By applying the Illumina Global

8 Screening Array (GSA) to 264 DNA samples with systematically altered quantity and quality,

9 we empirically tested the impact of SNP microarray analysis of deprecated DNA on kinship

10 classification success, as relevant in IGG. Reference data from manufacturer-recommended

11 input DNA quality and quantity were used to estimate genotype accuracy in the compromised

12 DNA samples and for simulating data of different degree relatives. Although stepwise decrease

13 of input DNA amount from 200 nanogram to 6.25 picogram led to decreased SNP call rates

14 and increased genotyping errors, kinship classification success did not decrease down to 250

15 picogram for siblings and 1st cousins, 1 nanogram for 2nd cousins, while at 25 picogram and

16 below kinship classification success was zero. Stepwise decrease of input DNA quality via

17 increased DNA fragmentation resulted in the decrease of genotyping accuracy as well as

18 kinship classification success, which went down to zero at the average DNA fragment size of

19 150 base pairs. Combining decreased DNA quantity and quality in mock casework and skeletal

20 samples further highlighted possibilities and limitations. Overall, GSA analysis achieved

21 maximal kinship classification success from 800-200 times lower input DNA quantities than

22   manufacturer-recommended, although DNA quality plays a key role too, while compromised

23   DNA produced false negative kinship classifications rather than false positive ones.

24

Author Summary

26   Investigative genetic genealogy (IGG), i.e., identifying unknown perpetrators of crime via

27   genomic database-tracing of their relatives by means of microarray-based single nucleotide

28   polymorphism (SNP) data, is a recently emerging field. However, SNP microarrays were

29   developed for much higher DNA quantity and quality than typically available from crime

30   scenes, while SNP microarray data on quality and quantity compromised DNA are largely

31   missing. As first attempt to investigate how SNP microarray analysis of quantity and quality

32   compromised DNA impacts kinship classification success in the context of IGG, we performed

33   systematic SNP microarray analyses on DNA samples below the manufacturer-recommended

34   quantity and quality as well as on mock casework samples and on skeletal remains. In addition

35   to IGG, our results are also relevant for any SNP microarray analysis of compromised DNA,

36   such as for the DNA prediction of appearance and biogeographic ancestry in forensics and

37   anthropology and for other purposes.

Introduction

For almost three decades, forensic DNA profiling with standard sets of polymorphic short tandem repeat (STR) markers has successfully been used to identify perpetrators of crime, thereby contributing towards solving numerous criminal cases worldwide (1, 2). However, in principle, forensic STR profiling is unsuitable for identifying unknown perpetrators, whose STR profiles are not yet included in national forensic DNA databases or are unknown to the investigative authorities otherwise. This consequently leads to cold cases with STR profiles of crime scene stains being available but not matching any known suspect, including all criminal offenders stored in the national forensic DNA database. Such situation allows perpetrators to continue their criminal activities and justice is denied to victims of crime or their families.

Over the last years, several DNA-based approaches for tracing unknown perpetrators have emerged. One such way is to search for family members of the unknown stain donor based on the forensic STR-profiles in the national forensic DNA database (i.e., familial search) (3). This approach, however, is limited in its success to first degree relatives, because of the limited number of autosomal STRs used in routine forensic DNA profiling (4). This disadvantage can be overcome by applying male-specific STRs from non-recombining regions of the Y-chromosome (Y-STRs) that can highlight male relatives of the unknown male perpetrator, given that the vast majority of perpetrators of major crimes are males (5). However, because forensic DNA databases in almost all countries consist of autosomal STR profiles, but not Y-STR profiles, Y-STR based familial search is restricted to voluntary DNA mass screenings (5). A more indirect way to find unknown perpetrators via focused police investigation is through investigative leads obtained via the prediction of externally visible characteristics of the unknown stain donor from crime scene DNA, including appearance traits (6), bio-geographic ancestry (7), and chronological age (8), in the context of Forensic DNA Phenotyping (9). Most

62 recently, investigative genetic genealogy (IGG) has started to emerge as new approach to find

63 unknown perpetrators with the help of DNA (4, 10).

64   IGG, also known as forensic genetic genealogy (FGG), is based on genomic data from

65 hundreds of thousands of autosomal single nucleotide polymorphisms (SNPs) typically

66 generated with SNP microarray technology (10). Because of the large number of autosomal

67 SNPs involved, IGG allows close and distant relatives from both, maternal and paternal sides

68 to be identified (4, 11). Over the last years, genomic databases consisting of high-density SNP

69 data have emerged, albeit outside the forensic field. Instead of governmental authorities, these

70 genetic genealogy databases are managed by private companies or private citizens, such as

71 GEDmatch consisting of around 1.3 million high-density SNP profiles as of 2020 (12). The

72 most commonly used method for identifying relatives in IGG is via DNA segments shared

73 between the unknown perpetrator obtained from crime scene DNA and individuals in genomic

74 databases (10, 13, 14). These identical-by-descend (IBD) segments that originate from the

75 same ancestor signal a familial relation between the highlighted person in the genomic database

76 and the unknown perpetrator, and the length of the shared segments translates into how close

77 the family relationship is. Police investigation to find the unknown perpetrator is then focused

78 on the highlighted relative via genealogical research (10).

79   The recent resolution of the Golden State Killer case in the USA, together with several

80 other cases, demonstrated the power of IGG (15, 16). As of November 2020, IGG had assisted

81 in over 200 cold cases, of which at least 28 were solved (17) (18). In May 2019, GEDmatch,

82 the most prominent genomic database used for IGG, updated its privacy regulations to require

83 users to opt-in for law enforcement to search their SNP profiles, and by October 2019, only

84 ~185,000 GEDmatch users had done this (12), reflecting a dramatic decrease of law

85 enforcement access compared to previous times when GEDmatch access for law enforcement

86 was unrestricted. In December 2019, the GEDmatch genomic database was acquired by the

87 forensic genomics firm Verogen. The company FamilyTreeDNA (FTDNA) also hosts a SNP

88 microarray database of 2 million of its customers that was initially established for other than

89 forensic reasons, while the company actively works with law enforcement to permit database

90 access for IGG (12, 19). FTDNA customers now need to opt-out to restrict law enforcement

91 from using their SNP profiles for searches. Currently, all genomic databases available for IGG

92 consist of SNP microarray data obtained from DNA of customer collected cheek swab or saliva

93 samples, also known as direct-to-consumer genetic testing (10). In principle, SNP profiles

94 extracted from whole genome sequencing (WGS) data obtained from crime scene DNA can

95 also be used to search DTC SNP databases given the SNP overlap, which was recently

96 exemplified in a murder case in Sweden (20). However, genomic databases for IGG that are

97 based on WGS data are yet to be established.

98   Besides the various ethical, societal, regulatory and other dimensions in relation to the

99 forensic use of genomic databases that were not established for forensic purposes (4, 10, 21),

100 there is an important technical dimension related to SNP microarray typing of forensic DNA

101 such as for IGG purposes. Notably, all currently available SNP microarrays were developed

102 and optimized for relatively high input DNA quality and quantity, which typically is not

103 available from human biological stains found at crime scenes. Moreover, the DNA

104 hybridization principle underlying all SNP microarray technologies is not expected to be well-

105 suited for compromised input DNA of low quantity and quality typically available from crime

106 scene stains. While the statistical methods to genetically classify relatives from high-density

107 SNP data in the context of IGG are being established throughout the last years (14, 22), studies

108 that applied SNP microarrays to compromised DNA samples are scarce (23-25). More

109 importantly, as far as we are aware, in-depth studies to systematically test for the impact the

110   use of SNP microarrays in compromised DNA has on kinship classification success in the

111   context of IGG are missing completely from the scientific literature as of yet. It has been

112   recognized by several authors that the lack of empirical data on the performance of SNP

113   microarrays in compromised DNA and its consequence on forensic use marks a serious

114   problem for the forensic application of SNP microarrays in general and IGG in particular (4,

115   10, 26).

116   In this study, we performed systematic SNP microarray experiments on hundreds of

117   DNA samples from multiple individuals with varying degrees of DNA input quantity and

118   quality to test the impact of resulting SNP microarray genotyping errors on kinship

119   classification success in the context of IGG. From the individual DNA samples of which we

120   used DNA samples of varying quality and quantity for SNP microarray analysis, we also

121   generated high-quality reference data from the high quality and quantity DNA input conditions

122   recommended by the microarray manufacturer. Together, these data were used to calculate

123   genotype accuracy for the compromised DNA samples. The high-quality data were additionally

124   used to simulate data of different degree relatives for kinship classification. To get further

125   insights, we additionally performed SNP microarray genotyping on DNA samples with a

126   combined decrease of quantity and quality i.e., DNA from mock casework samples of varying

127   storage temperature, time, and artificial degradation as well as naturally degraded DNA

128   obtained from skeletal remains. The overall SNP microarray data set we generated in this study

129   allowed us to quantify the effect of decreased DNA quality and quantity on SNP microarray

130   genotyping quality with regard to the ability to classify relatives of different degrees of

131   relationship. This data is vital for the development of IGG applications based on SNP

132   microarray analysis of DNA obtained from human crime scene samples in forensic genetic

133   casework, or from low quality and/or quantity DNA samples for other purposes in forensic

134   casework and anthropological studies.

135

136   **Results**

137   To evaluate SNP microarray technology for quantity and/or quality compromised input DNA

138   and its consequence for kinship classification in the context of IGG, we performed a series of

139   SNP microarray experiments with the widely used Illumina Global Screening Array (GSA).

140   We started out by generating microarray data of 24 individuals by using high input DNA

141   quantity and quality by meeting the recommendations of the microarray manufacturer i.e., 200

142   ng of high molecular weight DNA. This reference dataset was then used to conditionally

143   simulate SNP data of relatives of these 24 individuals based on four degrees of relationships

144   i.e., full siblings, 1st cousins, 2nd cousins, and 3rd cousins (**Figure S1**). We also used this high-

145   quality reference dataset to determine genotyping errors in the data obtained from the

146   compromised DNA samples of the same individuals, respectively (see methods for details).

147

148   *Impact of decreased input DNA quantity on SNP microarray-based kinship classification*

149   We quantified the impact of decreased input DNA quantity on the success of kinship

150   classification based on SNP microarray data. For this, we generated 8 GSA datasets per each

151   of 24 individuals based on manufacturer-recommended 200 ng input DNA as well as based on

152   seven lower amounts i.e., 1000 pg, 250 pg, 100 pg, 50 pg, 25 pg, 12.5 pg, and 6.25 pg, and

153   used them to perform kinship classification.

154   We found that on average the success of kinship classification remained over 98.5%

155   for input DNA amounts from 200 ng down to 250 pg for siblings and 1st cousins, and down to

156   1ng for 2nd cousins (**Figure 1**). With 25 pg (approximately 4 cells worth of DNA) and lower

157   input DNA amounts, the kinship classification success was zero for relatives of all four degrees

158   (**Figure 1, Table S1**). We observed considerable variation between DNA samples of the same

159   input amount between the 24 individuals on how decreasing input DNA quantity impacted on

160   kinship classification success. Classification of siblings remained 100% correct for all 24

161   individuals from 200 ng down to 250 pg and for 17 individuals also with 100 pg (average

162   classification rate of 93%). With 50 pg, 5 individuals still had 100% success rate for sibling

163   classification, while with 25 pg 22 individuals showed 0% classification success for siblings

164   and any other relatives tested. For 1st cousins, 100% classification success was revealed for all

165   24 individuals with 200 ng and 1 ng, while with 250 pg it was for 22 individuals (average

166   classification success of 99.4%), and with 100 pg for 13 individuals (76.6% success on average

167   across individuals). None of the 24 individuals had 100% classification success when applying

168   50 pg input DNA (18.5% success on average). For 2nd cousins, classification success was close

169   to 99% for all 24 individuals with 200 ng, which remained stable down to 1 ng (98.6% success

170   on average) and dropped to 91.6% with 250 pg, 55.7% with 100 pg and 6.6% with 50 pg. For

171   3rd cousins, we first note that even with ideal input DNA quantity of the manufacturer-

172   recommended 200 ng, classification success was on average only 74.6% across the 24

173   individuals, with no individual sample having more than 81.5% success rate, likely due to

174   insufficient identical-by-descent (IBD) sharing. The classification success rate was slightly

175   reduced from 1 ng input DNA (72.6% success). Using 250 pg input DNA, the 3rd cousin

176   classification success dropped down to 62.6%, with all but one sample decreasing in

177   classification success by more than 1%, and further down to 30.7% with 100 pg, where all

178   samples had decreased success rate but none yet at zero, and down to 3.3% with 50 pg, where

179   11 samples had 0% classification success. From input DNA amounts of 25 pg and lower, 3rd

180   cousin classification success was zero for all 24 individuals. When the true kinship relations

181   were no longer classified correctly due to compromised input DNA quantity, misclassifications

182   always occurred at the lower degrees of relatedness (e.g., siblings misclassified as cousins) and

183   did not result in false overestimations of the classified degree of relationship.

184
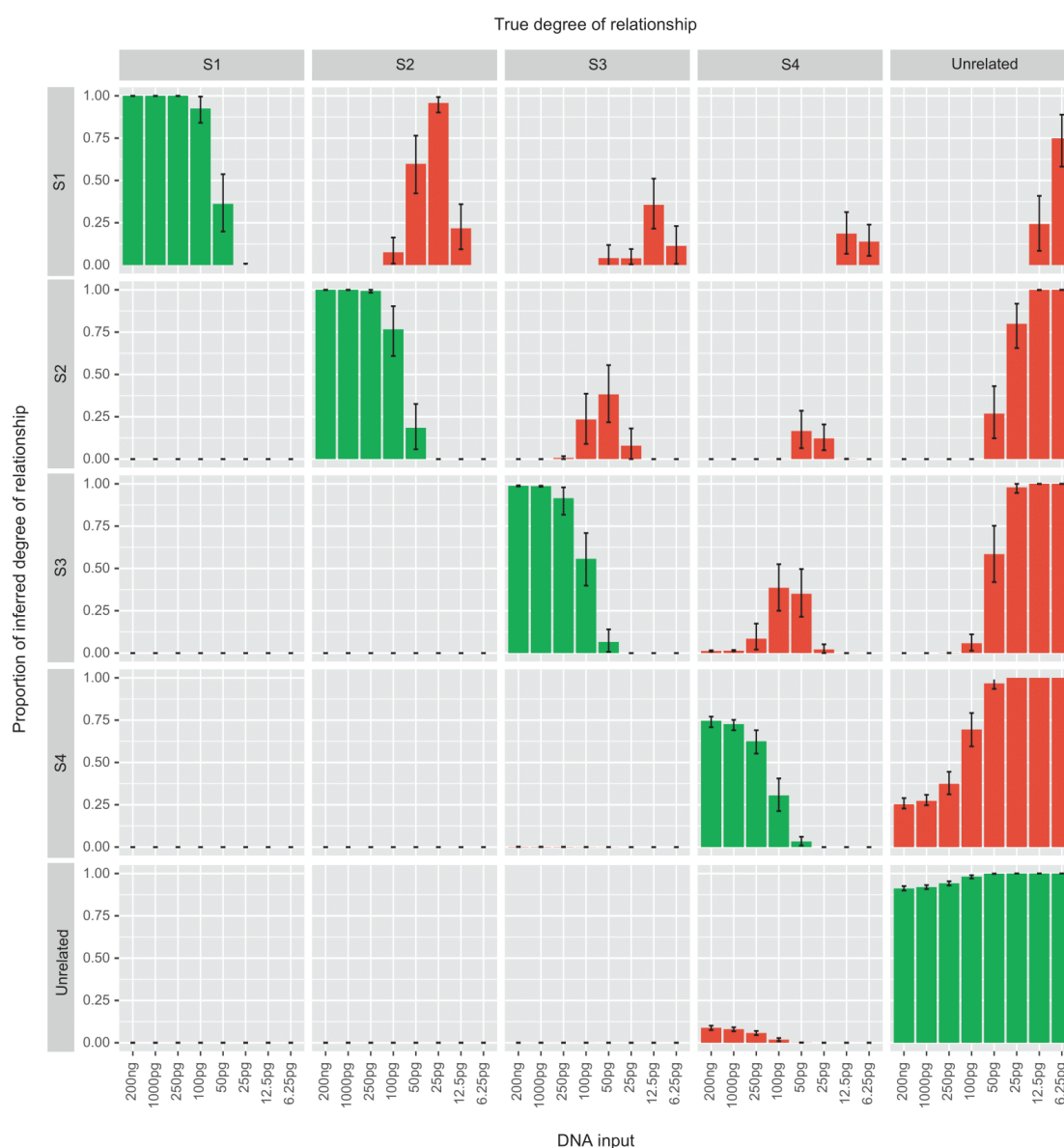
185   *Figure 1: Kinship classification of simulated relatives based on SNP microarray genotype data*

186   *obtained from input DNA of decreased quantity. Each plot in this matrix represents a simulated*

187   *kinship relation and a specific kinship classification based on a SNP microarray dataset obtained from*

188   *stepwise decreasing input DNA quantity. Each of the 24 individual DNA samples used in this*

189   *experiment had 1000 relatives simulated for each degree of relative, which were S1 (Full siblings), S2*

190   *(1st cousins), S3 (2nd cousins), S4 (3rd cousins) or unrelated from the high-quality reference data*

191   *generated from manufacturer-recommended input-DNA quantity and quality; the kinship classification*

192   *was restricted to these possible relation categories. For each of these 24 individuals, 8 input DNA*

193   *quantity levels i.e., 200 ng, 1 ng, 250 pg, 100 pg, 50 pg, 25 pg, 12.5 pg, and 6.25 pg were processed on*

194   *the GSA. The estimated genetic relation between such dataset from quantity-compromised DNA and*

195   *the relatives simulated from high-quality reference data was then classified per each of the 24*

196   *individuals. The x-axis of the matrix describes the classified kinship relations, while the y-axis of the*

197   *matrix describes the original simulated relations, across the 24 individuals used. For the individual*

198   *plots, the x-axis describes the input DNA quantity, and the y-axis describes the average kinship*

199   *classification rate. The green bar plots represent correctly classified relation and the red bar plots*

200   *represent the incorrect classifications. Error bars display the 75% confidence interval.*

201

Effect of decreased input DNA quantity on SNP microarray genotype accuracy

203   Aiming to better understand the observed impact of decreased input DNA quantity on

204   decreased kinship classification success, we investigated the effect of decreased DNA quantity

205   on SNP microarray genotype accuracy by calculating microarray-based SNP call rates and

206   genotype errors rates depending on varying input DNA amounts (as detailed in the methods).

207   We found that the stepwise decrease of input DNA quantity from optimal 200 ng down

208   to 6.25pg led to a gradual decrease in the SNP call rate (**Figure 2, Table S2**). In particular,

209   with manufacturer-recommended 200 ng, call rates for all 24 individuals were high as expected

210   at 99.9% (SD 0.0%), and decreased to an average of 96.7% (SD 2.5%) with 1 ng, 90.6% (SD

211   5.0%) with 250 pg, 84.4% (SD 4.5%) with 100 pg, and further down to a minimal call rate of

212   42.8% (SD 6.3%) from the lowest amount of 6.25 pg. While the number of called SNPs

213   decreased with decreasing input DNA amounts, an increase in the number of false genotypes

13

214   was seen, as may be expected, while differently so for the different types of genotype errors.

215   Errors were classified as "false heterozygotes" when a homozygous SNP was incorrectly typed

216   as a heterozygote, and "false homozygotes" when a heterozygous SNP was incorrectly typed

217   as a homozygote. The error rate started to increase already with the first DNA dilution step

218   below the manufacturer recommended 200 ng i.e. with 1 ng, albeit with a small average of

219   1.8% (SD 2.4%) for heterozygote errors, while homozygote errors remained close to zero at

220   0.01% (SD 0.01%) (**Figure 2, Table S2**). With 250 pg the heterozygote error rate increased to

221   an average of 7.9% (SD 6.8%) and the homozygote error rate to an average of 0.15% (SD

222   0.27%). Both error rates gradually increased further with further stepwise decreased input DNA

223   amounts, with the heterozygote error rate more so than the homozygote error rate. With 25 pg,

224   where the kinship classification success rate was zero for all degrees of relationship, the

225   heterozygote error rate was on average 43.3% (SD 9.7%), while the homozygote error rate was

226   8.1% (SD 2.6%). With the lowest input DNA amount of 6.25 pg, the heterozygous and

227   homozygous error rates were highest with on average 72.6% (SD 2.7%) and 21.3% (SD 4.3%),

228   respectively (**Figure 2, Table S2**).

229
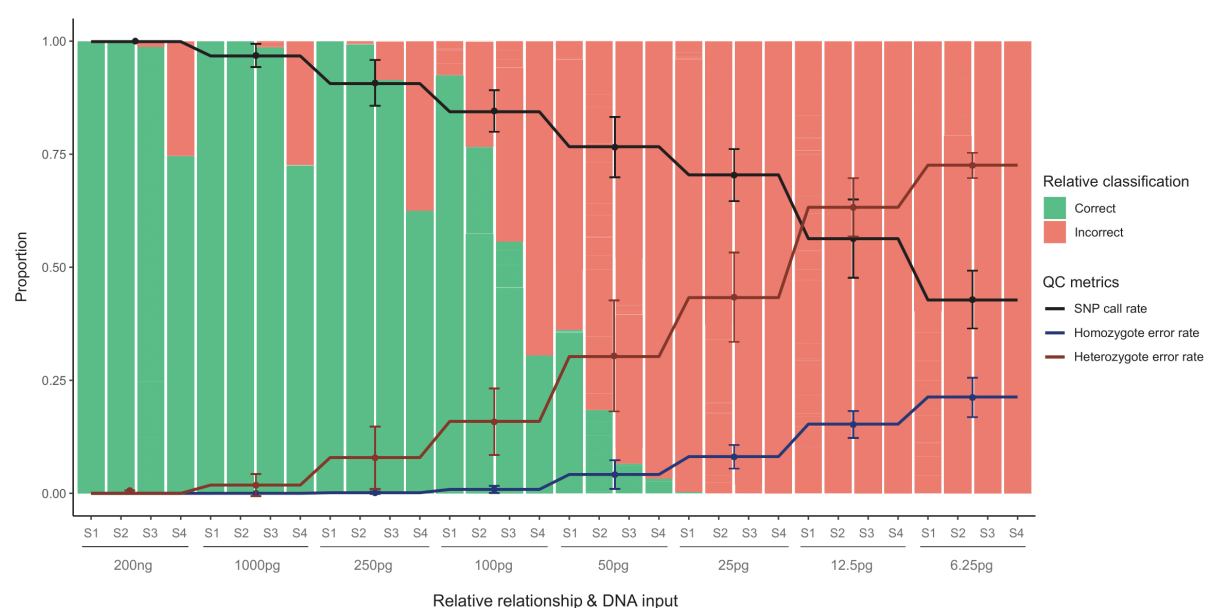
230   *Figure 2: Quality metrics of SNP microarray data from input DNA of decreased quantity.*

231   *The x-axis represents the 8 stepwise decreased input DNA quantities tested (200 ng – 6.25 pg)*

232   *and the degrees of kinship classification (S1 - Full siblings, S2 - 1ˢᵗ cousins, S3 - 2ⁿᵈ cousins,*

233   *S4 - 3ʳᵈ cousins) across the 24 individuals used. The y-axis runs from 0-100%, depicting call*

234   *rate, homozygote error rate, heterozygote error rate and classification success. Genotype call*

235   *rate and error rates are the average of the 24 genotype datasets obtained for each input DNA*

236   *quantity level. Error bars represent the 75% confidence interval.*

237

238   ***Impact of decreased input DNA quality on SNP microarray-based kinship classification***

239   Further, we investigated the effect of compromised DNA quality on kinship classification

240   success. This included non-degraded DNA samples as recommended by the microarray

241   manufacturer and degraded samples in three DNA fragmentation steps with decreased average

242   fragment size of 1000 bp, 500 bp and 150 bp for 12 individuals, of which 1 ng of DNA was

243   used for SNP microarray analysis (details in the methods section). Data from sample #526 at

244   the fragmentation level of 1000 bp was removed from the final data set for being an outlier

245   (Supplementary Table 9).

246   We observed that a stepwise decrease of input DNA quality by increasing the severity

247   of DNA fragmentation led to a gradual decrease in kinship classification success (**Figure 3,**

248   **Table S3**). More specifically, while an average fragment size of 1,000 bp (least severe

249   degradation level tested) had no effect on the classification success for siblings for all 12

250   individuals (100% success), the first cousin classification success of two individuals was no

251   longer 100% (average success of 99.8%). A more severe effect was seen for the classification

252   of 2ⁿᵈ cousins, where success decreased to an average of 89.7% from the 98.3% obtained with

253   non-fragmented DNA. An average fragment size of 500 bp resulted in a significant drop of

254   kinship classification success for all degrees of relatives tested i.e., to 56.7% for siblings (5

15

255   samples (41.7%) still at 100% success), 45.8% for 1st cousins (1 sample still at 100% success,

256   while 4 samples were at 0% success), and 19% for 2nd cousins (5 samples at 0% success rate).

257   As also seen in the DNA quantification experiments for optimal DNA amounts, for 3rd cousins

258   the classification success was already reduced with non-degraded DNA (average 70.8%), and

259   decreased further to 53.7% at 1000 bp fragmentation, and further to 9.5% at 500 bp. Finally,

260   the genotype profiles obtained from the most severely fragmented DNA samples of 150 bp

261   were insufficient to perform any accurate kinship classification (0%) for any of the relatives

262   tested. Similar to the DNA quantification experiments, there were no cases where compromised

263   input DNA quality resulted in an increase in the degree of kinship misclassification (no false

264   positives); in fact, all samples not correctly classified only saw a decrease in the degree of

265   kinship misclassification (**Figure 3, Table S3**).

266
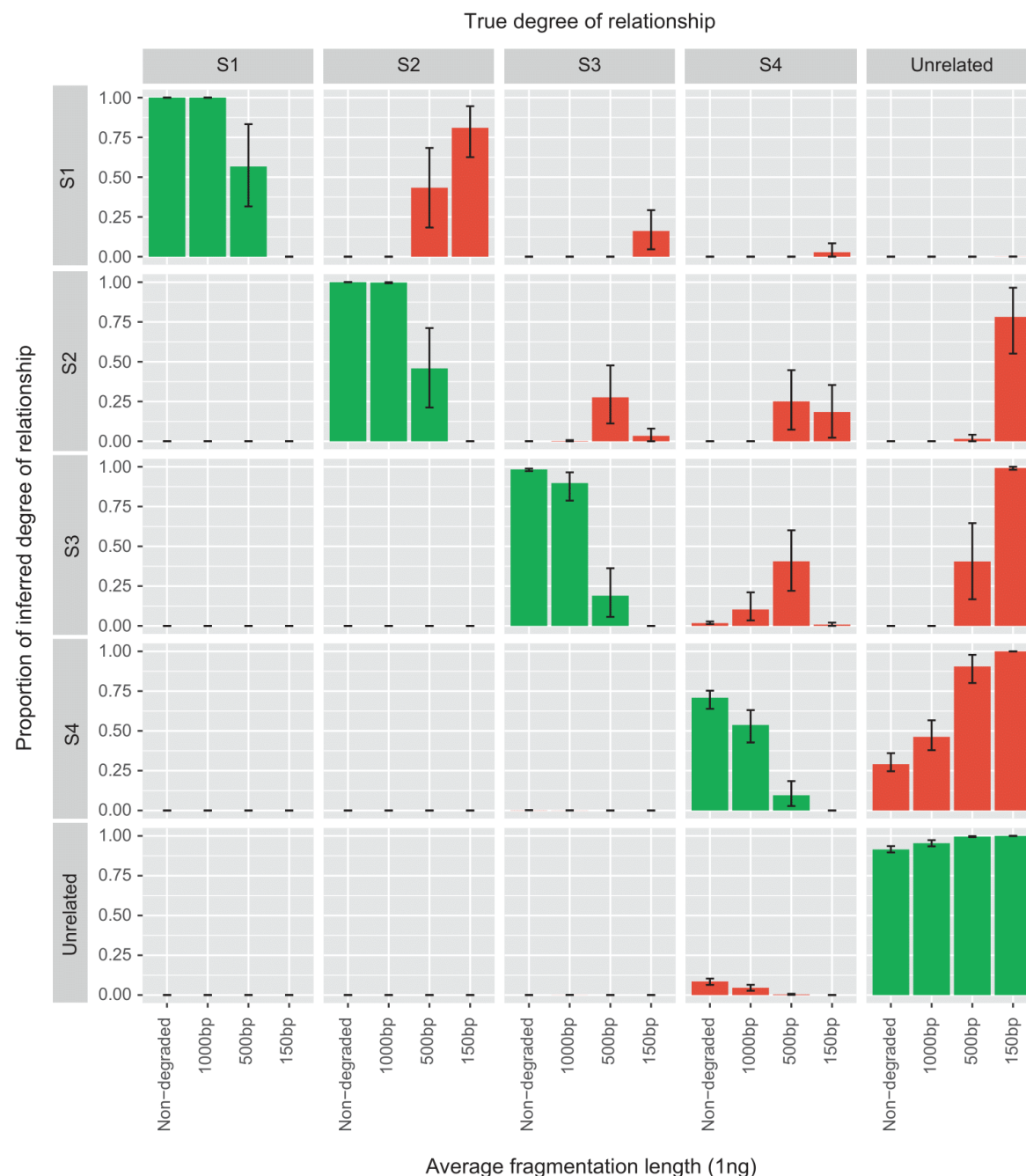
Figure 3: Kinship classification of simulated relatives based on SNP microarray data from input DNA of decreased quality. Each plot in this matrix represents a simulated kinship relation and a specific kinship classification based on a SNP microarray dataset obtained from decreasing input DNA fragment size. Each of the 12 samples used in this experiment had 1000 relatives simulated for each degree of relative, which were S1 (Full siblings), S2 (1ˢᵗ cousins), S3 (2ⁿᵈ cousins), S4 (3ʳᵈ cousins) or unrelated from the high-quality reference data generated from manufacturer-recommended input-DNA quantity and quality; the kinship classification was also restricted to these possible relation categories. For each of these 12 individuals, four input DNA fragmentation levels i.e., unfragmented, 100 bp, 500

17

276   *bp, 150 bp were processed on the GSA to obtain a genetic dataset. The estimated genetic relation*

277   *between such quality-compromised dataset and the relatives simulated from high quality reference data*

278   *was then classified for each of the 12 individuals. The x-axis of the matrix describes the classified*

279   *relations, while the y-axis of the matrix describes the original simulated relations, across the 12*

280   *individuals used. For the individual plots, the x-axis describes the average fragment size of the input*

281   *DNA, and the y-axis describes the average kinship classification rate. The green bar plots represent*

282   *correctly classified relation and the red bar plots represent the incorrect classifications. Error bars*

283   *display the 75% confidence interval.*

284

285   ***Effect of decreased input DNA quality on SNP microarray genotype accuracy***

286   Aiming to better understand the observed negative impact decreased DNA quality has on

287   kinship classification success, we studied the effect of quality-compromised input DNA on

288   SNP call rate and genotype error rate. Stepwise decrease of DNA quality by increase of DNA

289   fragmentation from non-fragmented down to highly fragmented resulted in a gradual decrease

290   in SNP call cate and an increase in genotyping errors, as may be expected. In particular, an

291   average fragment size of 1,000 bp had an immediate negative effect on the SNP call rate with

292   an average of 84.6% (SD 4.4%) and a negative effect on genotype accuracy with increasing

293   heterozygous error of 16.6% (SD 6.5%) on average and homozygous error of 0.2% (SD 0.2%),

294   compared to the non-degraded DNA samples with averages of 96.5% (SD 1.7%) 1.7% (SD

295   1.3%) and 0.0% (SD 0.0%), respectively. An average fragment size of 500 bp resulted in a

296   further decrease of the SNP call rate to an average of 75.1% (SD 5.7%), and a further decrease

297   of the genotype accuracy with increasing average heterozygote error of 32.4% (SD 10.0%) and

298   2.3% (SD 1.5%) homozygote error. Finally, the most severely fragmented input DNA of 150

299   bp had an average SNP call rate of 62.7% (SD 7.2%), an average heterozygote error of 53.6%

300   (SD 9.7%), and an average homozygote error of 8.1% (SD 2.7%) (**Figure 4, Table S4**).

301
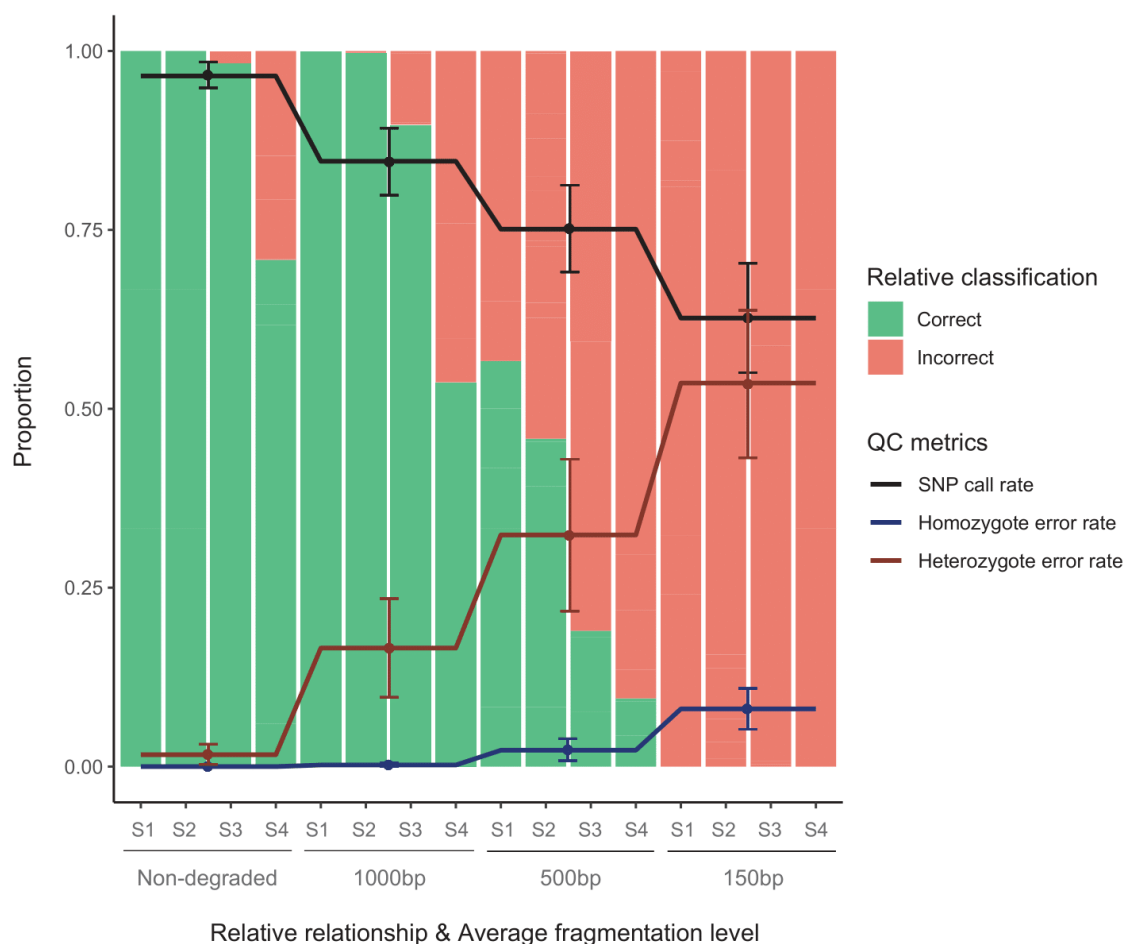
Figure 4: Quality metrics of the GSA genotype data from input DNA of decreased quality. The x-axis represents the 4 stepwise fragmentation levels tested (unfragmented, 1000bp, 500bp and 150bp) and the degrees of kinship classification (S1 - Full siblings, S2 - 1st cousins, S3 - 2nd cousins, S4 - 3rd cousins) across the 12 individuals used. The y-axis runs from 0-100%, depicting call rate, homozygote error rate, heterozygote error rate and classification success. Genotype call rate and error rates are the average of the 12 genotype datasets obtained for each input DNA fragmentation level. Error bars represent the 75% confidence interval.

Additional factors influencing SNP microarray genotype accuracy and SNP microarray-based kinship classification

To further test SNP microarray performance on the type of DNA samples typically confronted with in forensic casework, and its impact on kinship classification success in the context of

19

315   IGG, we performed SNP microarray analysis of forensic mock casework DNA samples. Mock

316   casework samples are a typical element of forensic validation studies; they are produced in a

317   way to mimic real crime scene samples. These mock casework samples were produced from

318   blood of individuals for which we had generated high-quality reference data, thus allowing to

319   estimate genotype accuracy and testing the impact on kinship classification. To this end, we

320   ran a total of 55 mock casework DNA samples from i) one individual's whole blood in making

321   24 bloodstains of different size (blood volume), prepared on different substrates, and exposed

322   under different environmental conditions such as storage time, temperature, humidity, and UV

323   radiation, the latter to mimic sun exposure, ii) blood DNA samples from three individuals at

324   200 ng on eight different artificial PCR inhibition levels, and iii) seven non-human DNA

325   samples (for details see material and methods) (**Table S5**).

326   Overall, we found that varying levels of blood stain storage conditions regarding

327   humidity, substrate type and storage time (up to 25 days) did not seem to have any effect on

328   the SNP microarray genotype accuracy and thus not on the kinship classification success

329   (**Table S5**). In contrast, and as expected, the size of bloodstain (DNA quantity) and DNA

330   damage via direct UV treatment (DNA quality) both appeared to affect genotype accuracy and

331   kinship classification success (**Table S5**). Particularly for the smallest bloodstains produced

332   from 1 µl of blood, the total isolated DNA amounts ranged from 0.167 to 5.163 ng depending

333   on various conditions. The effect of low DNA amount was evident in bloodstain 17 (296 pg),

334   which resulted in a lower accuracy of classification of 3rd cousins (47.1%), in line with what

335   we expected from the DNA quantification experiments (**Figure 1**, **Table S1**). However, when

336   we combined the effect of low DNA quantity with low DNA quality by damaging small

337   amounts of DNA with direct UV treatment (30 minutes) as in bloodstain 22 (167 pg), we

338   observed a decreased classification accuracy for 2nd cousins of 52.2% compared to 89.7% in

339   the non-treated DNA sample (**Table S5**). Notably, this decrease of kinship classification

340   success was not evident in bloodstain 23 despite the increased time of UV exposure (60

341   minutes), likely caused by the higher input DNA amount (1.35 ng) in this sample (**Table S5**).

342   Finally, our chosen PCR inhibitor, hematin (27), did not seem to affect SNP microarray

343   genotype accuracy. Independently from the hematin concentration, all samples yielded perfect

344   call rates of over 99% (**Table S6**). Given the very high call rate observed, estimation of kinship

345   classification success was not performed with the SNP array data from the hematin experiment.

346   Moreover, SNP microarray genotyping of all seven animal cell DNA samples, derived from

347   common pets and domesticated animals whose DNA is often recovered from crime scenes,

348   resulted in very low SNP call rates below 70% (**Table S7**), highlighting the human specificity

349   of the GSA. Notably, such low call rates were in our systematic DNA quantity and quality

350   experiments only observed for the 'worst' human DNA samples, namely those with lowest

351   input DNA quantity (12.5 and 6.25 pg) (**Figure 2**) and quality (150 bp fragment size) (**Figure**

352   **4**).

353

354   *Correlation between QC metrics from SNP microarray data and SNP microarray-based*

355   *kinship classification*

356   In typical practical applications of SNP microarray genotyping, such as for IGG but also for

357   other forensic and non-forensic purposes, only the SNP call rate is available as quality control

358   (QC) measure. Nevertheless, users of SNP microarray technology in practical applications

359   shall be interested in the reliability of the data they obtained from their analyzed DNA samples.

360   As a first step towards providing guidance on this matter of high practical relevance, we tested

361   if our diverse set of experimentally generated SNP microarray data from various input DNA

362   quantities and qualities would allow us to establish a preliminary measure on the reliability of

21

363   SNP microarray when deviating from the manufacturer recommendations for input DNA

364   quantity and quality. To this end, aiming at maximizing the statistical power, we pooled all

365   data we generated from high and low quality and quantity human DNA samples i.e., from DNA

366   quantity, DNA quality, and mock human bloodstain experiments, resulting in a combined

367   dataset containing call rate, genotype errors rates and classification successes from a total of

368   264 DNA samples **(Table S5, Table S8, Table S9)**. Using these data, and by employing n-

369   parameter logistic regression to fit a model, we correlated total SNP call rate with i) total

370   genotype error rate, ii) heterozygote error rate and iii) homozygote error rate and obtained very

371   high positive r^2 estimates of 0.98, 0.99, and 0.95, respectively (**Figure 5a**). Furthermore, we

372   correlated total SNP call rate with kinship classification success rate for all four degrees of

373   relatives (**Figure 5b**), where the fitted model achieved an r^2 of 0.89 for siblings, 0.87 for 1st

374   cousins, 0.95 for 2nd cousins, and 0.94 for 3rd cousins. We also made a model correlating

375   homozygote error rate with kinship classification success and achieved an r^2 of 1.00 for

376   siblings, 1.00 for 1st cousins, 1.00 for 2nd cousins, and 0.98 for 3rd cousins (**Figure S2**).

377

22

378
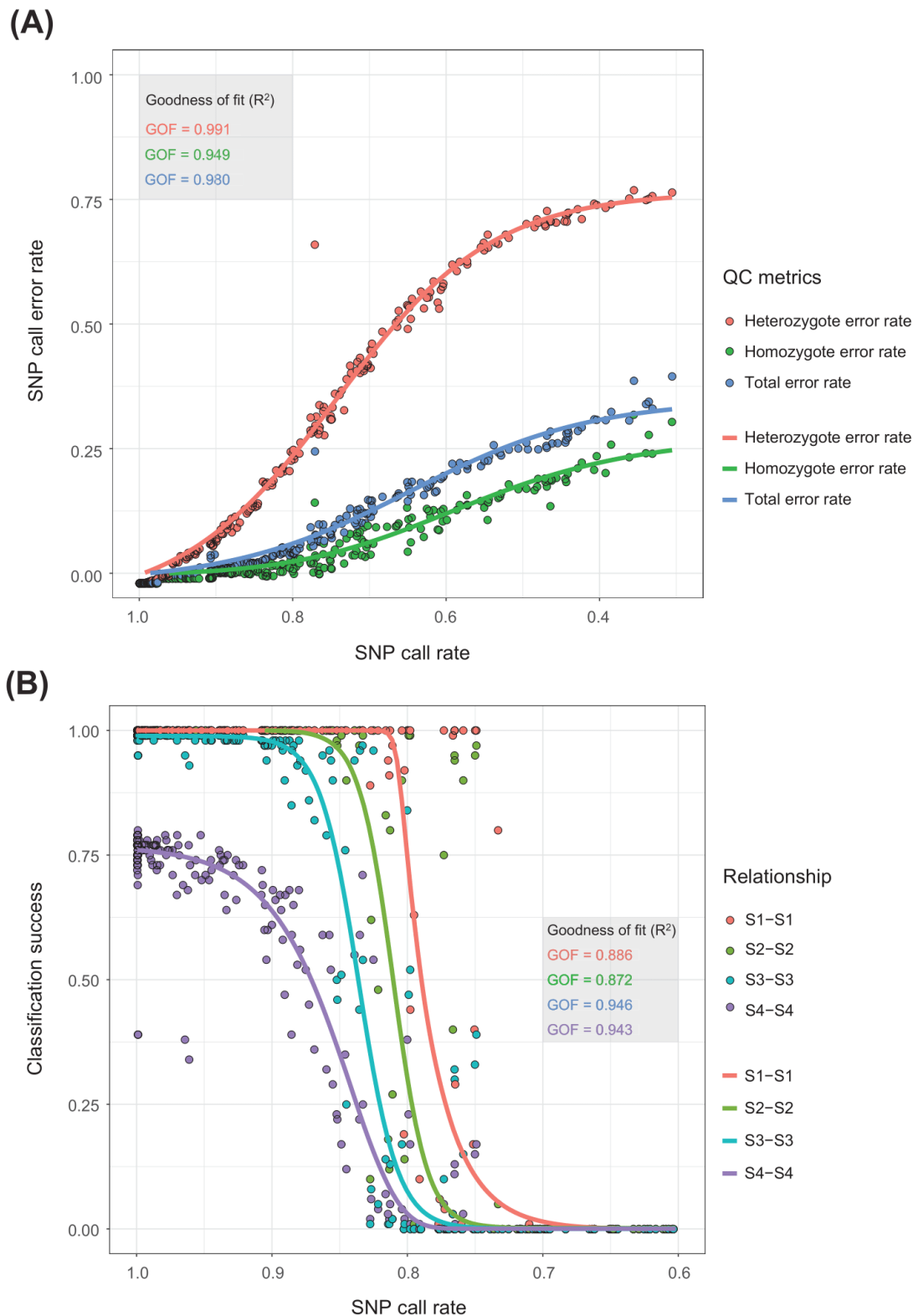
*Figure 5: **Relations of SNP microarray data quality metrics and kinship classification**. To describe*

the relation between the studied quality control metrics of our experiment, models were fitted with n-

parameter logistic regression, using the genotype error, call rate and kinship classification success

23

382 *data from the compromised quantity and quality as well as mock casework SNP microarray analysis in*

383 *a total of 264 samples (Table S5, Table S8, Table S9). **A)** Relationship of SNP call rate and genotyping*

384 *error rate. Three types of genotyping errors were considered for this model: homozygote error,*

385 *heterozygote error, and the total genotyping error by combining both. The homozygote error model had*

386 *a standard error (SE) of 0.016 with a goodness-of-fit (GoF / r^2) of 0.948, the heterozygote error model*

387 *had a SE of 0.0247 with a GoF of 0.991, and the total error model a SE of 0.014 with a GoF of 0.980.*

388 ***B)** Relationship of SNP call rate and kinship classification success. One model was fitted for each*

389 *degree of relation: S1-S1 (full sibling), S2-S2 (1st cousin), S3-S3 (2nd cousin), and S4-S4 (3rd cousin).*

390 *The full sibling classification success prediction model had a SE of 0.163 with a GoF of 0.886. The 1st*

391 *cousin success prediction model had a SE of 0.174 with a GoF of 0.872. The 2nd cousin success*

392 *prediction model had a SE of 0.110 with a GoF of 0.946. The 3rd cousin success prediction model had*

393 *a SE of 0.084 with a GoF of 0.943.*

394

395 As may be expected, we saw that all three genotyping error rates (total, heterozygote

396 and homozygote) increased as the SNP call rate decreased, with the heterozygote error being

397 impacted the most (**Figure 5a**). While the heterozygote error rate increased immediately with

398 decreased SNP call rate, the homozygote error rate remained below 0.05% as long as 93% of

399 the array SNPs were called (**Figure 5a**). More importantly, we found that within certain

400 boundaries it was possible to predict kinship classification success solely based on SNP call

401 rate (**Figure 5b**). Whereas for the most distant relationships, i.e., 3rd cousins, kinship

402 classification rate gradually decreased with call rate, the success of classifying siblings, as well

403 as 1st and 2nd cousins, remained at 100% until the call rate reached certain thresholds. These

404 thresholds were different for the different degrees of relationship and were lower the more

405 closely related the family members were. With SNP call rates below these thresholds, kinship

406 classification success decreased rapidly (**Figure 5b**). More specifically, we observed a

407 remaining high kinship classification success until, and drastic decrease below, 90% SNP call

408 rate for 2nd cousins, ~86% for 1st cousins, and ~82% for siblings. Universally and independently

409 of family relationship, we found that DNA samples with SNP call rates below 75% appeared

410   unsuitable for correctly inferring kinship relations, as below this threshold we observed 0%

411   classification success for all four degrees of relationship.

412

413   *SNP microarray analysis of naturally degraded DNA from skeletal remains*

414   Finally, to assess SNP microarray performance on naturally degraded DNA, thereby combining

415   compromised DNA quantity with compromised DNA quality through a natural body

416   decomposition process, we analyzed DNA samples obtained from skeletal remains buried

417   underground or found on forest surface that had experienced various natural conditions for

418   various times of postmortem intervals. A total of 18 bone or teeth derived DNA samples were

419   GSA-genotyped (**Table S10**). These DNA samples were additionally genotyped with a

420   commercial forensic STR kit commonly used for forensic identification. The latter allows a

421   comparison between the performance of both DNA technologies. Due to the nature of this

422   sampling set-up, SNP microarray reference data obtained from manufacturer-recommended

423   input DNA quantity and quality were not available for the individuals of which bone DNA

424   samples were analyzed, thus not allowing genotype accuracy estimations and kinship

425   classification success quantifications. However, by applying our established relation between

426   SNP call rate and kinship classification success (see above, **Figure 5b**), we used the SNP call

427   rate obtained from these naturally degraded DNA samples as proxy for SNP microarray data

428   quality, and from the obtained results, provided expectations on kinship classification success

429   in case such samples would be used for IGG purposes.

430   Two out of the 18 bone DNA samples (bones 17 & 18) with the highest recovered DNA

431   amounts of 27.09 and 78.5 ng fully consumed for GSA analysis, produced microarray-based

432   SNP genotype profiles with a call rate above 95% (**Table S10**). According to our fitted models,

25

433 these data could be used to correctly classify relatives up to 3rd cousins (**Figure 5b**). These two

434 bone DNA samples also yielded complete 24-loci forensic DNA profiles. Two other bone DNA

435 samples (bones 14 & 16) with total input DNA amounts of 2.16 ng and 3.14 ng yielded call

436 rates of 85.3% and 88.7%, respectively (**Table S10**), which according to our fitted models

437 could still be used to classify relatives up to 3rd cousins (**Figure 5b**), albeit with reduced success

438 rate. Sample 14 yielded a complete 24-locus forensic DNA profile, while sample 16 missed

439 two non-autosomal loci. Additionally, 12 of the bone DNA samples with total input DNA

440 amounts at the picogram level resulted in call rates below 70% (ranging from 32.6% to 69.7%),

441 which according to our models is unsuited to obtain any accurate kinship classification (**Figure**

442 **5a**). Of these 12 bone DNA samples, 7 gave complete 24-loci forensic DNA profiles, while 5

443 had partial profiles between 14 and 22 of the 24 markers. The remaining two bone DNA

444 samples (bones 1 and 2) with total input DNA amounts of 36 pg and 56 pg failed the GSA

445 analysis completely as they produced no signals. In the forensic DNA profiling, these two DNA

446 samples had produced results for 16 and 22 of the 24 loci, respectively.

447

448

449 **Discussion**

450 Our work represents the first study that systematically explores the impact SNP microarray

451 analysis of quantity and quality compromised DNA has for kinship classification success,

452 which is relevant for IGG. Earlier studies on SNP microarray genotyping of forensic DNA

453 samples did not report genotype reliability and whether the genetic data would be applicable

454 for kinship classification (18) or investigated the use of SNP microarrays for low quantities of

455 DNA in less samples and not analyzing DNA below 1 ng (23, 25). Moreover, as far as we are

456   aware, no previous study empirically investigated the impact of quantity and/or quality

457   compromised DNA on SNP microarray-based kinship classification success, which, however,

458   is highly relevant for applying SNP microarrays for IGG purposes in forensic practice.

459   The model commonly used in genetic genealogy for classifying relatives from genomic

460   data implements versions of the so-called segment approach (13, 28-30), which we also applied

461   here in a modified way. The novelty in our approach lies in the combination of conditional

462   simulation with the final step in which the original (perfect), gold standard reference genotype,

463   is replaced with an imperfect sample (e.g., DNA of low quantity and/or quality). Using this

464   novel approach on the SNP microarray data acquired from optimal manufacturer-

465   recommended input DNA of high quantity and quality, 25% of the 3rd cousins were not

466   correctly classified, while for 2nd cousins this only was 1%, and for 1st cousins and for siblings

467   we had 100% classification success (**Figure 1**). This finding is in line with expectations, as the

468   more distant the family relationship is, the less likely one will be able to detect shared DNA

469   because recombination can make the shared segment shorter or can lead to not inheriting the

470   shared DNA at all (31).

471   In our quantity-compromised DNA experiments, an amount of 250 pg of high

472   molecular weight DNA yielded similarly high success rates of near 100% for classifying

473   siblings and 1st cousins, and 1 ng for 2nd cousin, as did the manufacturer-recommended 200 ng.

474   Although these reduced input DNA quantities resulted in decreased SNP call rates and

475   increased rates with which homozygous loci are erroneously called as heterozygotes, we did

476   not notice any measurable impact on kinship classification success. This picture changed when

477   the input DNA quantity was decreased below these critical amounts, which resulted in

478   gradually decreased kinship classification success down to 25 pg DNA and lower, where

479   kinship classification success was zero for all four degrees of relatives tested. While an increase

480   in the homozygote error rate was observed, the kinship classification success rates dropped.

481   On the other hand, in our quality-compromised DNA microarray experiments, when 1

482   ng of DNA was gradually fragmented and used for SNP microarray analysis, the first

483   fragmentation level of 1000 bp yielded similarly high success rates of 100% for classifying

484   siblings and 1st cousins as the non-fragmented input DNA did, while decreases in SNP call

485   rates and increases in genotype error rate were seen compared to non-fragmented input DNA.

486   For 2nd cousins and 1000 bp fragmentation, a reduced classification success was found, which

487   further decreased for all relatives with further increased fragmentation levels. At the highest

488   fragmentation level of 150 bp, the kinship classification success was zero for all degrees of

489   relatives. These findings together imply the main deciding factor if kinship classification based

490   on SNP microarray data from compromised DNA is successful or not is the amount of proper

491   quality DNA in a sample.

492   Another novelty of our work is that we empirically determined how GSA microarray

493   genotype data is affected when compromised DNA is analyzed. We found that false

494   heterozygotes were much more prevalent than false homozygotes when degrading the DNA

495   quality and quantity. Previous work has highlighted how detrimental genotyping errors are for

496   relative classification (32) (33). We analyzed both genotyping error types separately as these

497   were different in prevalence and how they affected relative classification rate. Our data strongly

498   suggests that the false homozygote genotypes are the driving cause for relative

499   misclassification. This is expected, as the segment approach is highly sensitive to homozygote

500   genotype errors. Only a false homozygote can cause IBS0 (where neither of the alleles are

501   similar between the individuals) and pre-maturely end a shared segment. This is corroborated

502   by our finding that homozygote genotyping error rate and kinship classification success rates

503   are strongly correlated (r^2 of 0.999 for siblings, r^2 of 1.00 for 1st cousins, r^2 of 0.999 for 2nd

504   cousins and a r^2 of 0.976 for 3rd cousins) (**Figure S2**). The false homozygote genotype calls

505   will terminate segments in the kinship classification model rather than prolonging them; thus,

506   resulting in a decreased degree of inferred relatedness instead of an increased one. This could

507   explain why decreasing DNA quantity or quality did not inflate the false positive kinship

508   classification rate i.e., the degree of relatedness was always underestimated in our experiments.

509   Other models to infer family relationships from genomic data may have a different

510   sensitivity to genotyping errors, which may be explored in future studies. In particular, the

511   likelihood ratio-based model, adopted in forensic STR profiling, could be an alternative (32,

512   34, 35). With this model, fewer markers are needed in the calculations compared to the segment

513   model (32). Nevertheless, a virtue of the segment approach used here is that it is unaffected by

514   the most frequent error of Illumina GSA SNP genotyping microarrays as we showed here i.e.,

515   heterozygote genotyping errors. Another advantage of the segment approach is that it allows

516   kinship classification to be estimated in the absence of specifically stated pairs of hypotheses

517   of relatedness, which instead is a key requirement for applying the likelihood ratio approach.

518   We note that several implementations of the segment approach mitigate issues with genotyping

519   errors, for instance through allowing a number of opposite homozygote genotypes in a shared

520   segment (29, 36).

521   As a side aspect, we studied the isolated effect of decreasing SNP numbers, as proxy

522   for SNP call rate, on the kinship classification success under optimal input DNA conditions.

523   For this, the used the high-quality reference data, leaving aside the effects of decreased input

524   DNA quantity and quality. This exercise illustrates that for closer relationships (e.g., siblings)

525   SNP genotyping call rates as low as 5.6% can still allow accurately classification through

526   detection of shared segments, while for more distant relationships (e.g. 3rd cousins) it needs

527   SNP genotyping call rates above 50% to be sufficient (**Figure S3**). These minimal SNP

528   numbers rather serve as theoretical expectations, whereas practically relevant thresholds of

529   SNP call rates to achieve trustworthy kinship classifications are seen in Figure 5. These

530   findings confirm our observations from quantity and quality compromised DNA samples that

531   SNP genotyping call rate alone is not driving kinship classification success (**Figures 2, 4 & 5**).

532   This further corroborates the finding that SNP genotyping errors are the main driving force of

533   kinship misclassification, even so, a decrease in SNP genotyping call rate can be accounted for

534   in the segment approach by tuning the number of SNPs required to define IBD segments.

535   However, we demonstrated that decreasing this threshold can have a detrimental impact on

536   kinship classification success (**Figure S4)**. Decreasing the number of SNPs per segment will

537   lead to false inclusion of random segments as IBD, thus, resulting in an estimated higher degree

538   of family relationship compared to the true degree.

539   Remarkably, we had fairly high kinship classification success with low input DNA

540   amounts of 100 pg. We attribute this to the whole genome amplification (WGA) step that is

541   part of the standard SNP microarray protocol. WGA increases the amount of DNA several

542   thousand fold (37, 38), potentially compensating in part for a low input DNA amount. WGA

543   is typically not utilized in forensic DNA analysis, as it can cause STR profiling to result in

544   extraneous bands (39) and unequal amplification between different targeted loci (40). While

545   these concerns are relevant for forensic STR profiling, the detrimental effect of WGA on SNP

546   microarray genotypes is reported to be considerably lower (38, 41). WGA-based SNP

547   microarray genotype data were found to be highly concordant with the SNP microarray data

548   from the respective gDNA (42). An additional benefit of WGA is that it is unaffected by the

549   presence of the PCR-inhibitor hematin (**Table S6**). However, the assumed WGA-based

550   compensation effect of low input DNA quantity has limits, as our study clearly demonstrated.

30

551   By decreasing the DNA quantity and quality below the recommendations of the

552   microarray manufacturer, we also observed a notable rise of genotyping errors, where the

553   increase of false heterozygous genotypes was larger than that of false homozygote genotypes

554   (**Figures 2 & 4**). The intrinsic mechanisms of the Illumina GSA suggest an explanation why

555   heterozygote genotype calls are more prone to be false than homozygote genotype calls. We

556   provide a brief description of Infinium technique behind Illumina GSA arrays in **Figure S5**.

557   We hypothesize that the majority of false genotype calls observed with decreased input DNA

558   quantity and quality was caused by decreasing probe intensity due to less DNA being bound to

559   hybridization probes, and a subsequent effect on the signal processing and genotype

560   interpretation software **(Figures S6 & S6)**. When probe intensity decreases, the chance

561   increases that the SNP call shifts to a different genotyping cluster. The heterozygote cluster is

562   found closest to the lowest intensity values, which may cause homozygotes to be called as

563   heterozygotes, explaining the higher heterozygote error rate compared to homozygote error

564   rate, which we observed at increasingly compromised input DNA for both input DNA quantity

565   and quality.

566   Judging the suitability of a DNA sample for SNP microarray analysis is generally

567   difficult since a number of factors such as quantity, fragmentation, and DNA damage could

568   impact genotype accuracy. In typical practical applications of SNP microarrays to

569   compromised DNA, such as IGG, reference data from manufacturer-recommended input DNA

570   quantity and quality are unavailable so that genotyping errors cannot be determined. Therefore,

571   we used quality metrics that are available for every SNP microarray dataset, such as SNP call

572   rate, to address if a SNP microarray dataset is suitable for kinship classification, or not. Our

573   results show that SNP genotyping call rate is highly correlated with genotyping error rate and

574   with kinship classification success rate (**Figure 5a&b**). Therefore, we suggest that the SNP

575 call rate, being the most common SNP microarray genotyping quality metric available, can be

576 used as a primary indicator to estimate the reliability of Illumina GSA genotype datasets, and

577 thus its suitability for kinship classification such as in IGG. In our data analysis, we observed

578 a steady increase in false SNP genotypes in SNP microarray datasets with SNP call rates below

579 99% (**Figure 5b**). Different SNP genotyping call rate thresholds are applied by different users,

580 with 95% being one standard call rate cut-off (43, 44). We found that samples with SNP call

581 rates of 90% or above yielded reliable kinship classification; while below 90% we started to

582 observe a decrease in kinship classification success, first for 2nd cousins, then on lower call rate

583 levels also for 1st cousins and for siblings (**Figure 5b**). Third cousin classification started at a

584 75% success rate even at high quality and quantity input DNA and declined gradually with

585 decreasing call rates. Based on our data, for SNP genotyping call rates below 75% it is unlikely

586 that kinship classification will be successful for any degree of family relationship.

587 Consequently, there is a grey zone between 75% and 90% call rate where the kinship

588 classification success is still high for close relationships (e.g., siblings), but quickly drops for

589 distant relationships (e.g., 2nd cousins). Notably, these thresholds coincide with the increase in

590 homozygote error rate, starting at roughly 90% call rate as well (**Figure 5b**).

591   Since experimental DNA degradation is artificial and does not resemble natural

592 degradation processes, we additionally analyzed bone and tooth derived DNA samples from

593 skeletal remains. These DNA samples naturally degraded under varying conditions and

594 combine decreased DNA quantity and decreased quality, which is typical in missing person

595 cases (45). However, only a minority of the skeleton-derived naturally degraded DNA samples

596 yielded high-enough call rates sufficient for concluding high genotype accuracy (two samples

597 had call rate > 93%, two others were in the grey zone >85% call rate). These four bone DNA

598 samples also had comparatively high amounts of quantified DNA, highlighting the importance

599   of input DNA amount for GSA analysis. The qPCR-measured input DNA amounts of these

600   four best performing bone DNA samples were all above 2.1 ng, while we obtained similar call

601   rates of ~85% using 100pg of non-fragmented DNA. As our DNA quality experiments have

602   shown that severely degraded DNA cannot be genotyped accurately with the GSA, the lower

603   SNP microarray genotyping reliability can be attributed to the poor quality of nuclear DNA in

604   the majority of the bone or teeth DNA samples tested (45). Notably, the SNP microarray results

605   of these bone and teeth derived DNA samples were not in good agreement with results from

606   forensic STR profiling. Although we found that 3 of the 4 samples with high SNP call rates

607   had complete forensic STR profiles, this was also seen for 7 of the samples with low SNP call

608   rates. A difference in success rates of forensic STR profiling and SNP microarrays is not

609   unexpected given the substantial differences in the underlying genotyping technologies and in

610   the number of analyzed DNA markers with each technology.

611   While the majority of naturally degraded DNA samples from skeletal remains did not

612   result in reliable SNP genotype profiles, our mock casework samples that simulated crime

613   scene samples generally did result in high quality genotyping profiles as well as high kinship

614   classification success (**Table S5**). This is most likely due to the fact that blood stains generally

615   are a good source for DNA (46) and that the artificial DNA degradation applied to the mock

616   casework samples was less severe in impact and time than the natural DNA degradation in the

617   skeletal remains. These preliminary data provide optimistic expectations that SNP microarray

618   genotyping might find application in the forensic cases, including IGG, as long as DNA yield

619   is high enough e.g., with input DNA amounts above 250-500 pg, and DNA degradation is

620   rather low, such as obtained from crime scene stains that experienced rather mild

621   environmental conditions similar to what we simulated in our mock casework samples. For

622   DNA samples of lower quantities and/or more severe degradation, as seen in the skeletal

623 remain samples and the severely fragmented DNA samples analyzed here, SNP microarray

624 analysis is rather not suitable. Besides DNA quantification to select samples for SNP

625 microarray analysis, the achieved SNP call rate and the preliminary models we introduced here

626 can be used as guidance on the suitability of the generated data such as for kinship classification

627 in IGG applications.

628 While our study is the first to empirically investigated the impact of quantity and/or

629 quality compromised DNA on SNP microarray-based kinship classification success in the

630 context of IGG, others previously reported on the use of SNP microarrays on forensic DNA

631 samples, or other samples with quantity limitations. The company Parabon Nanolabs (USA)

632 reported results from about 250 forensic case samples genotyped on the CytoSNP-850k array

633 (18), a predecessor of the Illumina GSA array (47). Roughly half of the DNA samples analyzed

634 with SNP microarrays were reported with SNP call rates above 95%, while also about half had

635 more than 10 ng of input DNA used (18). Based on our analyses, those cases are expected to

636 find kinship matches in IGG database investigations using the segment approach as long as the

637 respective relative is included in the database used. Wendt et al. (23) recently performed DNA

638 titration experiments from 200 ng to 1 ng on 3 DNA samples obtained from postmortem blood

639 samples with Illumina array (Omni2.5Exome-8 v1.3) genotyping as part of their DNA input

640 pretesting for a genome-wide association study (GWAS). They found that the SNP microarray

641 genotypes from 200 ng down to 25 ng did not differ significantly and thus used 25 ng as input

642 DNA threshold for their GWAS. Similar to our findings, which were based on larger sample

643 size and larger titration span going down to 6.25 pg, the authors showed higher rates of allelic

644 drop-in (false heterozygotes) than allelic drop-out (false homozygotes) with decreasing DNA

645 quantity. Another recently published study explored how long-time stored serum samples, with

646 an average input DNA amount of 5.8 ng, performed on two other types of SNP microarrays,

34

647   the Affymetrix Axiom Array (Thermofisher) and the Illumina HumanCoreExome array. In

648   80% of these DNA samples, SNP call rates above 94% were reported. In comparison, 92% of

649   our 1 ng high-quality DNA samples had SNP call rates above 94%. The level of degradation

650   was not quantified in that previous study, which could account for the lower number of reported

651   samples with high call rate. However, the authors had made alterations to the standard SNP

652   microarray genotyping protocol to account for the expected low DNA quality, given the use of

653   long-time stored serum. For Axiom Array analysis the number of WGA cycles was doubled,

654   and for the Illumina Human Core Exome array analysis DNA restoration kits were used (25).

655   These are options that might improve SNP call rates, which we did not explore here and may

656   deserve systematic investigation in the future.

657   In conclusion, our study provides the first empirical evidence how SNP microarray

658   analysis of quantity- and quality-compromised DNA impacts on kinship classification success,

659   which is highly relevant in the context of IGG. The GSA used here as an example of a widely

660   used SNP microarray was able to obtain high-density SNP profiles that were accurate enough

661   to achieve high kinship classification success at 800 times lower input DNA amounts for

662   siblings and 1st cousins, and 200 times lower for 2nd cousins, compared to the amount

663   recommended by the microarray manufacturer. Decreasing DNA quantity or quality reduced

664   the number of called SNPs and introduced heterozygote and, to a lesser extent, homozygote

665   genotyping errors. These SNP microarray genotyping errors, believed to be caused by low

666   probe intensity, are the primary cause of incorrect kinship classification as result of using

667   compromised DNA for SNP microarray analysis. When applied for IGG purposes, the

668   consequence may potentially be to miss a match to a relative included in the genetic genealogy

669   database, while the genotyping errors are unlikely to result in a false positive kinship match as

670   our study implies. Overall, our results are relevant for SNP microarray applications of

671   compromised DNA for IGG purposes, aiming at finding unknown perpetrators of crime via

672   their relatives stored in genomic databases accessible by law enforcement agencies. They are

673   also relevant for other forensic SNP microarray applications such as appearance prediction and

674   biogeographic ancestry inference of unknown perpetrators or missing persons from

675   compromised DNA, aiming at helping to identify unknown perpetrators and missing persons

676   via focused investigative intelligence.

677

678

679   **Materials and Methods**

680   *Samples.* Two types of human biological samples were used in this study: i) high-quality

681   EDTA blood samples from 24 completely anonymous blood donors of European descent that

682   A.U.G. received in 2003 from the Sanquin blood bank in Rotterdam, Netherlands, for the

683   purpose of genetic research and genetic method evaluation, and ii) DNA samples from 18

684   skeletons collected by a forensic pathologist after medico-legal examination of human remains

685   found on a forest surface or exhumed for genetic identification, at the Department of Forensic

686   Medicine in Krakow, Poland, including teeth, skull bones, humeri, femoral, clavicle and

687   metacarpal bone. Sample details and description of DNA isolation and quality control can be

688   found in **Supplementary Methods**.

689   *DNA titration.* The manufacturer-recommended amount of input DNA for the Illumina GSA

690   microarray, used here as a prominent SNP microarray example, is 200 ng (48). This is beyond

691   what is typically available in forensic crime scene samples where the DNA yield is highly

692   various and depends on the size of the crime scene stain and other parameters. In this series of

693   experiments, we therefore tested for the effect of the input DNA amount below the optimal 200

694   ng on the SNP microarray performance in terms of genotyping suboptimal amounts of DNA

695   i.e. 1000 pg, 250 pg, 100 pg, 50 pg, 25 pg, 12.5 pg and 6.25 pg, produced via DNA titration

696   and confirmed via a qPCR assay (Qiagen Investigator Quantiplex Kit). The minimal amount

697   of 6.25 pg roughly equals the genomic DNA content of one human diploid cell. Given that for

698   each of the 24 individuals tested we analyzed 8 different input DNA amounts, this DNA

699   quantity experiment included a total of 192 DNA samples that were processed for SNP

700   microarray genotyping.

701   *DNA fragmentation.* The manufacturer's recommendation on input DNA quality for the

702   Illumina GSA microarray is to use high molecular weight DNA. DNA samples obtained from

703   crime scene stains are usually of low molecular weight due to DNA degradation caused by

704   various factors that impact on biological stains at the scene of crime, such as temperature and

705   humidity, or long time periods since sample deposition at the scene of crime prior to collection.

706   In this series of experiments, we therefore tested for the effect of input DNA with decreased

707   quality. We opted for DNA fragmentation as one component of DNA degradation and used the

708   adaptive focused acoustics technology (Covaris), which has the advantage of fragmenting

709   DNA to a preset average fragment length. For 12 samples, randomly sampled from the initial

710   24, we degraded 1 ng of DNA to average fragment sizes of 1,000 bp, 500 bp, and 150 bp as

711   analyzed on the Labchip GX (Perkin Elmer). Given that for each of the 12 individuals tested

712   we analyzed 4 different input DNA amounts, this DNA quality experiment included a total of

713   48 DNA samples that were processed for SNP microarray genotyping.

714   *Mock casework.* To test for the performance of the Illumina GSA microarray on forensic-type

715   samples, we generated mock casework samples that mimic crime scene stains. In total, 24

716   bloodstains from whole blood of one of the 24 blood donors were exposed to different

717   conditions that are often encountered in forensic samples. The factors we considered included

37

718   stain size (the volume of blood used), storage time, storage temperature, substrate (the surface

719   blood was deposited), relative humidity and DNA damage (via UV radiation), as detailed in

720   **Table S5**. The total amount of DNA extracted from each bloodstain was used as input in the

721   Illumina GSA array, ranging from 0.167 to 127.8 ng of DNA. Additionally, we tested the

722   effects of PCR inhibitor hematin and non-human DNA on GSA genotyping, detailed in the

723   supplementary methods.

724   *Microarray genotyping.* The study used the Infinium Global Screening Array (GSA) for all

725   SNP microarray genotyping (48). The choice was made partly based on its popularity in the

726   Direct-to-Consumer Genetics industry, but also due to its rich up-to-date SNP content. Array

727   genotyping was performed according to standard Illumina protocols (37). We decided to refrain

728   from additional sample or data treatment, e.g. the use of restoration kits on the DNA samples

729   or application of Hardy-Weinberg equilibrium on the SNP data. Microarray scan data was

730   converted to genotypes using Genome Studio 2.0, the standard software for processing Illumina

731   SNP microarrays. As this project consisted of samples of sub-optimal input DNA quantity and

732   quality, standard microarray QC protocols, such as filtering on SNP or sample call rates, were

733   not executed. We extracted a subset of markers based on the overlap with the 1000 Genomes

734   phase III reference panel (49), resulting in a total of 519,300 SNP markers. Genetic positions

735   in centiMorgan (cM) were extracted from Rutger's map or alternatively interpolated for

736   markers absent from the repository (50). For each gold standard sample, the genotypes were

737   phased to generate haplotypes employing the Eagle V2.4 algorithm (51) with the European

738   (CEU) individuals from the 1000 Genomes Project as reference data (49).

739   *Genotype quality assessment.* The high-density SNP profiles obtained from GSA analysis of

740   each tested condition was rated by two main measures of quality: SNP call rate and genotype

741   discordance to the gold standard, which was obtained from manufacturer-recommended input

742 DNA quantity and quality for each of the 24 individuals. SNP call rate is the fraction of the

743 total probes with genotype calls and it is the most common quality measure used in SNP

744 microarray analysis. The number of false genotypes in each dataset was determined by direct

745 comparison of the genetic profile from each tested condition to their corresponding gold

746 standard reference dataset obtained from optimal DNA conditions. Heterozygous error rate is

747 defined as the percentage of false heterozygotes among all heterozygotes, and homozygous

748 error rate is defined as the percentage of homozygotes that are either heterozygotes or opposite

749 homozygote in the gold standard reference data.

750 ***Conditional simulations of relatives.*** For each of the 24 individuals, we generated relatives

751 through a conditional simulation approach using the gold standard reference dataset obtained

752 from manufacturer-recommended input DNA quantity and quality. **Figure S1** provides an

753 overview of our algorithm to simulate family members based on our gold standard samples. In

754 brief, the approach uses phased genotype data from high quality samples and proceeds by

755 conditionally generating relatives for each sample. We restricted the algorithm to siblings, 1st

756 cousins, 2nd cousins and 3rd cousins (In the figures abbreviated as S1, S2, S3 and S4), spanning

757 a range of the most relevant degrees of relationships for IGG purposes. For each of the 24 gold

758 standard samples, we repeated this relative simulation process 1000 times for each type of

759 relation. In detail, the simulation process starts at the first SNP on a chromosome and initiates

760 by drawing alleles from the gold standard sample with probabilities equal to the identical by

761 descent (IBD) probabilities for each degree of relationship (14, 52). Moving along the

762 chromosome, the process either continues with the same IBD state as for the previous marker

763 or change IBD state with probabilities equal to the rate of recombination between the markers.

764 Ultimately, the process generates a complete high-density SNP dataset for the relative with a

765 mosaic of shared DNA with the gold standard sample. In contrast to approaches that employ

766   population allele frequencies and gene dropping to generate pedigree data, our approach i)

767   takes into account the known phased genotype of the original sample and ii) uses phased

768   haplotypes from the population to draw non-IBD alleles; thus, mitigating potential biases

769   caused by linkage disequilibrium. The novelty in our approach is the combination of

770   conditional simulation with the final step in which the original, gold standard, genotype is

771   replaced with an imperfect sample.

772   ***Inferring degree of relatedness.*** Genealogy assessment, referred here as kinship classification,

773   was performed by inferring the degree of relatedness of the imperfect sample to the simulated

774   relatives (see Figure S1). The term *imperfect* is used as an umbrella for inhibited, degraded or

775   quantity-reduced DNA samples. We used a version of what we call the segment approach to

776   infer degree of relatedness between individuals. The algorithm is described in for example (13,

777   29, 30) and detailed for forensic purposes in Kling et al. (14). The version implemented by

778   direct-to-consumer company Ancestry.com is also outlined in a white paper (28). This is a

779   method that identifies haplotypes between individuals that are identical by descent (IBD)

780   without requiring information on allele frequencies in the population. Briefly, the most naïve

781   version of the approach measures stretches of genotypes where at least one nucleotide is

782   identical at each base position between the pair of individuals (IBS1 or IBS2) to ultimately

783   infer IBD segments. The stretches are only terminated if opposite homozygous genotypes are

784   detected (IBS0). Furthermore, in order to accurately define a segment as IBD the algorithm

785   needs thresholds, firstly the length as measured in cM and secondly the total number of

786   overlapping SNPs in each segment. Both thresholds have the purpose of excluding non-IBD

787   segments. The length restriction was tuned to include short enough segments for distant

788   relatives and still large enough not to include unrelated individuals as relatives. We used 5 cM

789   as the length detection threshold, corroborated by previous studies (53) as well as our own

790   studies (**Figure S7**). We further explored the impact of the second threshold i.e., the number

791   of overlapping SNPs required in each segment. As the imperfect samples may contain several

792   locus dropouts, a default requirement of say 500 SNPs can result in missed segments.

793   Therefore, we explored 100-700 SNPs as thresholds, detailed in **Figure S6**. Finally, in order to

794   explore the isolated effect of decreasing number of markers, we thinned the data, resulting in

795   reduced marker subsets ranging from the original 519,300 down to 10,379 markers (**Figure**

796   **S3**). We proceeded to infer relatedness with each reduced set of markers using the methods

797   previously described.

798

807

References

808   1.   Jobling MA, Gill P. Encoded evidence: DNA in forensic analysis. Nat Rev Genet.
809   2004;5(10):739-51.
810   2.   Kayser M, de Knijff P. Improving human forensics through advances in genetics, genomics
811   and molecular biology. Nat Rev Genet. 2011;12(3):179-92.
812   3.   Bieber FR, Brenner CH, Lazer D. Finding Criminals Through DNA of Their Relatives.
813   Science. 2006;312(5778):1315.
814   4.   Ge J, Budowle B. Forensic investigation approaches of searching relatives in DNA
815   databases. Journal of Forensic Sciences. 2020;66(2):430-43.

816 5. Kayser M. Forensic use of Y-chromosome DNA: a general overview. Hum Genet.
817 2017;136(5):621-35.
818 6. Kayser M. Forensic DNA Phenotyping: Predicting human appearance from crime scene
819 material for investigative purposes. Forensic Sci Int Genet. 2015;18:33-48.
820 7. Phillips C. Forensic genetic analysis of bio-geographical ancestry. Forensic Science
821 International: Genetics. 2015;18:49-65.
822 8. Freire-Aradas A, Phillips C, Lareu MV. Forensic Individual Age Estimation with DNA - From
823 Initial Approaches to Methylation Tests. Forensic Sci Rev. 2017;29:121.
824 9. Schneider PM, Prainsack B, Kayser M. The Use of Forensic DNA Phenotyping in Predicting
825 Appearance and Biogeographic Ancestry. Dtsch Arztebl International. 2019;116(51-52):873-80.
826 10. Kling D, Phillips C, Kennett D, Tillmar A. Investigative genetic genealogy: Current methods,
827 knowledge and practice. Forensic Science International: Genetics. 2021;52:102474.
828 11. Erlich Y, Shor T, Pe'er I, Carmi S. Identity inference of genomic data using long-range familial
829 searches. Science. 2018;362(6415):690-4.
830 12. Murphy KHH. Your DNA Profile is Private? A Florida Judge Just Said Otherwise. The New
831 York Times. 2019.
832 13. Browning BL, Browning SR. A fast, powerful method for detecting identity by descent. Am J
833 Hum Genet. 2011;88(2):173-82.
834 14. Kling D, Tillmar A. Forensic genealogy—A comparison of methods to infer distant
835 relationships based on dense SNP data. Forensic Science International: Genetics. 2019;42:113-24.
836 15. Phillips C. The Golden State Killer investigation and the nascent field of forensic genealogy.
837 Forensic Science International: Genetics. 2018;36:186-8.
838 16. DeLisi M. Forensic epidemiology: Harnessing the power of public DNA sources to capture
839 career criminals. Forensic Science International. 2018;291:e20-e1.
840 17. Katsanis SH. Pedigrees and Perpetrators: Uses of DNA and Genealogy in Forensic
841 Investigations. Annual Review of Genomics and Human Genetics. 2020;21(1):535-64.
842 18. Greytak EM, Moore C, Armentrout SL. Genetic genealogy for cold case and active
843 investigations. Forensic Science International. 2019;299:103-13.
844 19. Ram N, Roberts JL. Forensic genealogy and the power of defaults. Nature Biotechnology.
845 2019;37(7):707-8.
846 20. Tillmar A, Sjölund P, Lundqvist B, Klippmark T, Älgenäs C, Green H. Whole-genome
847 sequencing of human remains to enable genealogy DNA database searches – A case report.
848 Forensic Science International: Genetics. 2020;46:102233.
849 21. Murphy E. Law and policy oversight of familial searches in recreational genealogy databases.
850 Forensic Science International. 2018;292:e5-e9.
851 22. Kling D, Welander J, Tillmar A, Skare Ø, Egeland T, Holmlund G. DNA microarray as a tool in
852 establishing genetic relatedness—Current status and future prospects. Forensic Science
853 International: Genetics. 2012;6(3):322-9.
854 23. Wendt FR, Rahikainen AL, King JL, Sajantila A, Budowle B. A genome-wide association
855 study of tramadol metabolism from post-mortem samples. Pharmacogenomics J. 2020;20(1):94-103.
856 24. Cho S, Yu HJ, Han J, Kim Y, Lee J, Lee SD. Forensic application of SNP-based
857 resequencing array for individual identification. Forensic Science International: Genetics. 2014;13:45-
858 52.
859 25. Rounge TB, Lauritzen M, Erlandsen SE, Langseth H, Holmen OL, Gislefoss RE. Ultralow
860 amounts of DNA from long-term archived serum samples produce quality genotypes. European
861 journal of human genetics : EJHG. 2020;28(4):521-4.
862 26. Ge J, Budowle B. How many familial relationship testing results could be wrong? PLOS
863 Genetics. 2020;16(8):e1008929.
864 27. Sidstedt M, Steffen CR, Kiesler KM, Vallone PM, Rådström P, Hedman J. The impact of
865 common PCR inhibitors on forensic MPS analysis. Forensic Science International: Genetics.
866 2019;40:182-91.
867 28. Ball CA, Barber MJ, Byrnes J, Carbonetto P, Chahine KG, Curtis RE, et al., editors.
868 AncestryDNA Matching White Paper Discovering genetic matches across a massive, expanding
869 genetic database2016.
870 29. Henn BM, Hon L, Macpherson JM, Eriksson N, Saxonov S, Pe'er I, et al. Cryptic distant
871 relatives are common in both isolated and cosmopolitan genetic samples. PLoS ONE.
872 2012;7(4):e34267.
873 30. Gusev A, Lowe JK, Stoffel M, Daly MJ, Altshuler D, Breslow JL, et al. Whole population,
874 genome-wide mapping of hidden relatedness. Genome Res. 2009;19(2):318-26.

875   31.   Edge MD, Coop G. Donnelly (1983) and the limits of genetic genealogy. Theoretical
876   Population Biology. 2020;133:23-4.
877   32.   Kling D. On the use of dense sets of SNP markers and their potential in relationship
878   inference. Forensic Science International: Genetics. 2019;39:19-31.
879   33.   Monroy Kuhn JM, Jakobsson M, Günther T. Estimating genetic kin relationships in prehistoric
880   populations. PLOS ONE. 2018;13(4):e0195491.
881   34.   Gjertson DW, Brenner CH, Baur MP, Carracedo A, Guidet F, Luque JA, et al. ISFG:
882   Recommendations on biostatistics in paternity testing. Forensic Sci Int Genet. 2007;1(3-4):223-31.
883   35.   Abecasis GR, Cherny SS, Cookson WO, Cardon LR. Merlin - Rapid analysis of dense genetic
884   maps using sparse gene flow trees. Nat Genet. 2002;30(1):97-101.
885   36.   Browning BL, Browning SR. Detecting identity by descent and estimating genotype error rates
886   in sequence data. Am J Hum Genet. 2013;93(5):840-51.
887   37.   Illumina, editor Infinium HTS Assay: Reference Guide2019.
888   38.   Illumina, editor Infinium LCG Quad Assay Protocol Guide2015.
889   39.   Barber AL, Foran DR. The utility of whole genome amplification for typing compromised
890   forensic samples. J Forensic Sci. 2006;51(6):1344-9.
891   40.   Findlay I, Ray P, Quirke P, Rutherford A, Lilford R. Diagnosis and preventing inherited
892   disease: Allelic drop-out and preferential amplification in single cells and human blastomeres:
893   implications for preimplantation diagnosis of sex and cystic fibrosis. Human Reproduction.
894   1995;10(6):1609-18.
895   41.   Dagnall CL, Morton LM, Hicks BD, Li S, Zhou W, Karlins E, et al. Successful use of whole
896   genome amplified DNA from multiple source types for high-density Illumina SNP microarrays. BMC
897   genomics. 2018;19(1):182-.
898   42.   Deleye L, Vander Plaetsen A-S, Weymaere J, Deforce D, Van Nieuwerburgh F. Short
899   Tandem Repeat analysis after Whole Genome Amplification of single B-lymphoblastoid cells.
900   Scientific reports. 2018;8(1):1255-.
901   43.   Reed E, Nunez S, Kulp D, Qian J, Reilly MP, Foulkes AS. A guide to genome-wide
902   association analysis and post-analytic interrogation. Statistics in medicine. 2015;34(28):3769-92.
903   44.   Psaty BM, O'Donnell CJ, Gudnason V, Lunetta KL, Folsom AR, Rotter JI, et al. Cohorts for
904   Heart and Aging Research in Genomic Epidemiology (CHARGE) Consortium: Design of prospective
905   meta-analyses of genome-wide association studies from 5 cohorts. Circulation Cardiovascular
906   genetics. 2009;2(1):73-80.
907   45.   Watherston J, McNevin D, Gahan ME, Bruce D, Ward J. Current and Emerging Tools for the
908   Recovery of Genetic Information From Post Mortem Samples: New Directions for Disaster Victim
909   Identification. Forensic Sci Int Genet. 2018;37:270-82.
910   46.   Mapes AA, Kloosterman AD, van Marion V, de Poot CJ. Knowledge on DNA Success Rates
911   to Optimize the DNA Analysis Process: From Crime Scene to Laboratory. Journal of Forensic
912   Sciences. 2016;61(4):1055-61.
913   47.   Illumina, editor Infinium CytoSNP-850K BeadChip: Reference Guide2019.
914   48.   Illumina, editor Infinium™ Global Screening Array-24 v3.0 BeadChip2016.
915   49.   Auton A, Abecasis GR, Altshuler DM, Durbin RM, Abecasis GR, Bentley DR, et al. A global
916   reference for human genetic variation. Nature. 2015;526(7571):68-74.
917   50.   Matise TC, Chen F, Chen W, De La Vega FM, Hansen M, He C, et al. A second-generation
918   combined linkage physical map of the human genome. Genome Res. 2007;17(12):1783-6.
919   51.   Loh P-R, Danecek P, Palamara PF, Fuchsberger C, A Reshef Y, K Finucane H, et al.
920   Reference-based phasing using the Haplotype Reference Consortium panel. Nat Genet.
921   2016;48(11):1443-8.
922   52.   Weir BS, Anderson AD, Hepler AB. Genetic relatedness analysis: modern data and new
923   challenges. Nat Rev Genet. 2006;7(10):771-80.
924   53.   Durand EY, Eriksson N, McLean CY. Reducing pervasive false-positive identical-by-descent
925   segments detected by large-scale pedigree analysis. Mol Biol Evol. 2014;31(8):2212-22.

926

927   **Supporting information**

928   Supplementary Methods

929   Supplementary Figures

930   Figure S1. Description of our novel approach to conditionally generate relatives

931   Figure S2**.** Homozygote error rate *versus* family classification success

932   Figure S3. Isolated effect of call rate on family classification success.

933   Figure S4. Effect of segment SNP number and total call rates on family classification

934   success.

935   Figure S5. Simplified technical background of Illumina GSA.

936   Figure S6. Possible explanation fot the high number of false heterozygotes in

937   compromised DNA genotyping.

938   Figure S7. Effect of segment length on family classification success.

939   Supplementary Tables

940   Table S1. Proportion of inferred relationship for all degrees of relationship of the

941   titration samples.

942   Table S2. SNP microarray QC metrics of the titration samples

943   Table S3. Proportion of inferred relationship for all degrees of relationship of the

944   fragmented samples

945   Table S4. SNP microarray QC metrics of the fragmented samples

946   Table S5. Overview of the mock casework samples including QC metrics (call and

947   genotype discordance rates) and family classification success

948   Table S6. Overview of artificially inhibited samples and their call rates

949   Table S7. Overview of animal DNA samples and their call rates

950   Table S8: Individual sample data for the quantity experiment

951   Table S9: Individual sample data of the quality (DNA fragmentation) experiment

952   Table S10. Information on the samples obtained from skeletal remains included in this

953   study