# A High-Throughput, Rapid, Automated and Nanoscale SNP Fingerprinting Workflow for Assessing Sample Quality, Integrity and Contamination for Use in a Biorepository

Luke Stewart, Bill Hunt
Standard BioTools Inc., South San Francisco, CA, United States

ISBER 2025 | Poster PI-08

## Introduction

Biorepositories provide access to high-quality, curated samples for basic and clinical research purposes. Sample degradation, misidentification and contamination are significant risks to the integrity of banked samples. Distribution of such samples can waste time and laboratory resources and negatively impact the integrity and reproducibility of research studies.

Standard procedures for sample identity and traceability have been employed by biorepositories for many years, including barcode labeling and LIMS tracking. Establishing DNA identity for each sample using a single-nucleotide polymorphism (SNP) fingerprinting method is listed in the *ISBER Best Practices: Recommendations for Repositories: Fifth Edition* as part of quality control for nucleic acids. Implementing such a DNA fingerprinting method in the biorepository workflow provides a direct association of sample molecular identity.

The Advanta™ Sample ID Genotyping Panel is a 96-SNP assay that generates a sample-specific genetic fingerprint of research samples at any point in the sample journey. Targeted SNPs, which include 80 in exonic regions to support population prediction, also support the assessment of sample quality and determination of chromosomal sex. In this study, SNP fingerprints were created using the Advanta Sample ID Genotyping Panel with the Biomark™ X9 System for High-Throughput Genomics. SNP fingerprint analysis indicates that the panel can be used to assign individual identity, detect sample cross contamination, assess sample quality and identify samples from the same individual.

Based on Standard BioTools™ microfluidics technology, the Sample ID workflow uses integrated fluidic circuits (IFCs) to precisely generate multiple datapoints per sample through concurrent, independent assay reactions at nanoliter volumes. In this poster, we demonstrate the utility of the Advanta Sample ID Genotyping Panel and the Biomark X9 System as a sample identity and traceability tool that can be easily implemented into a routine biorepository operation.

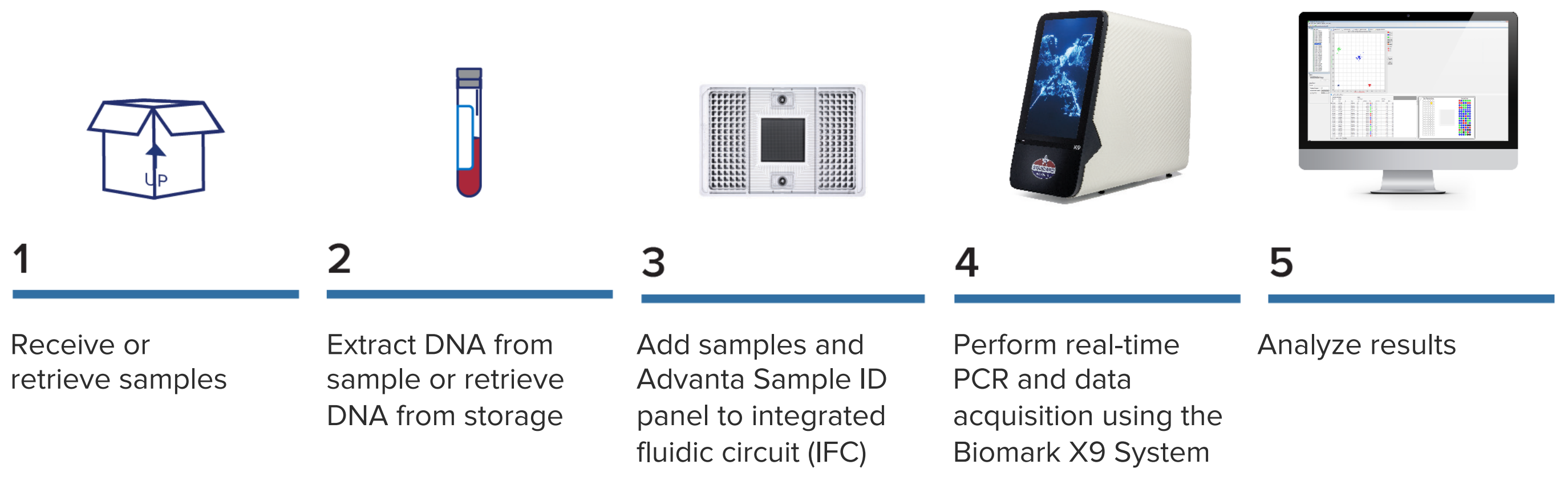
## Methods

The Advanta Sample ID Genotyping Panel coupled with the Biomark X9 System enabled automated genotype calling of 96 carefully selected SNPs in up to 96 individual samples simultaneously in approximately four hours. The 96.96 GT Preamp IFC-X was used, which enables processing of samples with as low as 2.5 ng/uL human genomic DNA.

The **Genotyping Using the 96.96 GT Preamp IFC-X with SNP Type Assays** protocol in the Biomark X9 System Gene Expression and Genotyping User Guide (FLDM-01040) was followed for all experiments conducted in this study. The Advanta Sample ID Genotyping Panel consists of:

- Ten quality SNPs located in regions susceptible to DNA degradation. Call rates correlate with sample quality.
- Six chromosomal sex SNPs where three are located in the X chromosome and three are located in the Y chromosome. Useful for identifying sample swaps and contamination.
- Forty population-specific SNPs located in exons within housekeeping genes. SNPs were selected with 0.5 minor allele frequency across three major HapMap populations. These SNPs provide high discriminatory power to differentiate individuals where a duplicate genotype probability is 1 in 1.09 x $10^{17}$.
- Forty highly polymorphic SNPs located in exons within housekeeping genes. SNPs were selected with 0.5 minor allele frequence in at least one of the three major HapMap populations to aid in population prediction.

The Biomark X9 System consists of a real-time qPCR instrument, IFCs and PCR reagents. The 96.96 GT Preamp IFC-X workflow using the Advanta Sample ID Genotyping Panel consists of five basic steps (Figure 1).



1 Receive or retrieve samples

2 Extract DNA from sample or retrieve DNA from storage

3 Add samples and Advanta Sample ID panel to integrated fluidic circuit (IFC)

4 Perform real-time PCR and data acquisition using the Biomark X9 System

5 Analyze results

**Figure 1. The Advanta Sample ID Genotyping Panel workflow on the Biomark X9 System, from sample preparation through data analysis.** Ninety-six samples and the 96-assay panel are loaded on an IFC and combined automatically in the Biomark X9 System where thermal cycling and imaging takes place. The genotyping run is then analyzed using Standard BioTools SNP Genotyping Analysis Software.

## Detecting sample cross contamination

Human genomic DNA from multiple research samples was diluted to 10 ng/uL in 1X DNA Suspension Buffer. Two sets of two samples were mixed by volume at ratios of 90:10, 50:50 and 10:90. Each sample mixture (including 100%) was run in duplicate in the IFC. Overall call rates and gender calls were calculated for each of the four samples and each mixture using Standard BioTools SNP Genotyping Analysis Software. Reduced overall call rates and no-calls for chromosomal sex correlated to cross-contaminated samples (Table 1). Increased heterozygote cluster spread was observed for cross-contaminated samples when compared with 100% samples (Figure 2).

| Sample Name | Overall Call Rate % | Autosomal No. of No-Calls | Autosomal SNP Performance Call | Chromosomal Sex Determination Call |
|---|---|---|---|---|
| FGM3 100% | 100.0 | 0 | Pass | Male |
| FGM3 100% | 100.0 | 0 | Pass | Male |
| FGM3 90%<br>FGM4 10% | 78.1 | 20 | Fail | No-Call |
| FGM3 90%<br>FGM4 10% | 77.1 | 20 | Fail | No-Call |
| FGM3 50%<br>FGM4 50% | 79.2 | 18 | Fail | No-Call |
| FGM3 50%<br>FGM4 50% | 75.0 | 22 | Fail | No-Call |
| FGM3 10%<br>FGM4 90% | 82.3 | 17 | Fail | No-Call |
| FGM3 10%<br>FGM4 90% | 82.3 | 17 | Fail | No-Call |
| FGM4 100% | 100.0 | 0 | Pass | Male |
| FGM4 100% | 100.0 | 0 | Pass | Male |

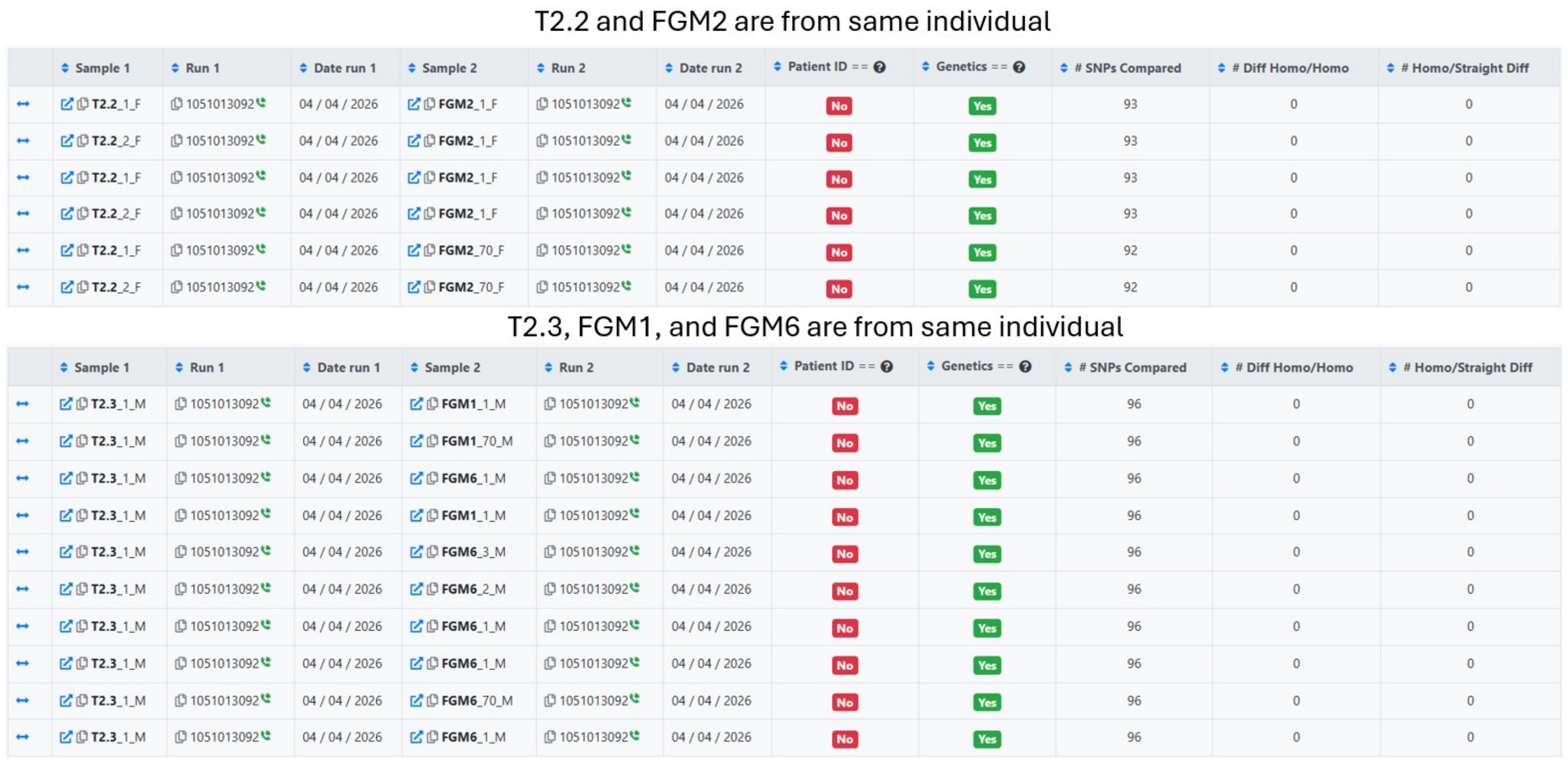
**Table 1. Cross-contamination samples and call rates to assess contamination.** Two male samples (FGM3 and FGM4) were mixed together at 90:10, 50:50 and 10:90 ratios by volume and run in duplicate. As contamination rate increases call rate decreases, and the samples are flagged for call performance. Additionally, contamination of 10% leads to no-calls for gender, further flagging the sample.

## Identifying multiple samples from the same individual

Multiple samples from the same individual were run in replicate at 10 ng/uL as previously described. Resultant data was generated using Standard BioTools SNP Genotyping Analysis Software, exported as a .csv file and subsequently imported into third-party software for assessment of genetic matching based on SNP fingerprints.

The third-party software compared SNP fingerprints across all individuals and samples and correctly determined that one individual was sampled three times, and another individual was sampled twice based on SNP call genetic similarity comparisons (Figure 3). SNP fingerprints of all technical replicates for each samples matched as expected.

Two separate father-mother-child trios were run to assess individual discriminatory power of the Advanta Sample ID Genotyping Panel. SNP fingerprints were compared within each trio and even though the individuals are related, their fingerprints were deemed sufficiently distinct to differentiate (Figure 5).



T2.2 and FGM2 are from same individual

| | Sample 1 | Run 1 | Date run 1 | Sample 2 | Run 2 | Date run 2 | Patient ID == | Genetics == | # SNPs Compared | # Diff Homo/Homo | # Homo/Straight Diff |
|---|---|---|---|---|---|---|---|---|---|---|---|
| ↔ | T2.2_1_F | 1051013092 | 04 / 04 / 2026 | FGM2_1_F | 1051013092 | 04 / 04 / 2026 | No | Yes | 93 | 0 | 0 |
| ↔ | T2.2_2_F | 1051013092 | 04 / 04 / 2026 | FGM2_1_F | 1051013092 | 04 / 04 / 2026 | No | Yes | 93 | 0 | 0 |
| ↔ | T2.2_1_F | 1051013092 | 04 / 04 / 2026 | FGM2_1_F | 1051013092 | 04 / 04 / 2026 | No | Yes | 93 | 0 | 0 |
| ↔ | T2.2_2_F | 1051013092 | 04 / 04 / 2026 | FGM2_1_F | 1051013092 | 04 / 04 / 2026 | No | Yes | 93 | 0 | 0 |
| ↔ | T2.2_1_F | 1051013092 | 04 / 04 / 2026 | FGM2_70_F | 1051013092 | 04 / 04 / 2026 | No | Yes | 92 | 0 | 0 |
| ↔ | T2.2_2_F | 1051013092 | 04 / 04 / 2026 | FGM2_70_F | 1051013092 | 04 / 04 / 2026 | No | Yes | 92 | 0 | 0 |

T2.3, FGM1, and FGM6 are from same individual

| | Sample 1 | Run 1 | Date run 1 | Sample 2 | Run 2 | Date run 2 | Patient ID == | Genetics == | # SNPs Compared | # Diff Homo/Homo | # Homo/Straight Diff |
|---|---|---|---|---|---|---|---|---|---|---|---|
| ↔ | T2.3_1_M | 1051013092 | 04 / 04 / 2026 | FGM1_1_M | 1051013092 | 04 / 04 / 2026 | No | Yes | 96 | 0 | 0 |
| ↔ | T2.3_1_M | 1051013092 | 04 / 04 / 2026 | FGM1_70_M | 1051013092 | 04 / 04 / 2026 | No | Yes | 96 | 0 | 0 |
| ↔ | T2.3_1_M | 1051013092 | 04 / 04 / 2026 | FGM6_1_M | 1051013092 | 04 / 04 / 2026 | No | Yes | 96 | 0 | 0 |
| ↔ | T2.3_1_M | 1051013092 | 04 / 04 / 2026 | FGM1_1_M | 1051013092 | 04 / 04 / 2026 | No | Yes | 96 | 0 | 0 |
| ↔ | T2.3_1_M | 1051013092 | 04 / 04 / 2026 | FGM6_3_M | 1051013092 | 04 / 04 / 2026 | No | Yes | 96 | 0 | 0 |
| ↔ | T2.3_1_M | 1051013092 | 04 / 04 / 2026 | FGM6_2_M | 1051013092 | 04 / 04 / 2026 | No | Yes | 96 | 0 | 0 |
| ↔ | T2.3_1_M | 1051013092 | 04 / 04 / 2026 | FGM6_1_M | 1051013092 | 04 / 04 / 2026 | No | Yes | 96 | 0 | 0 |
| ↔ | T2.3_1_M | 1051013092 | 04 / 04 / 2026 | FGM6_1_M | 1051013092 | 04 / 04 / 2026 | No | Yes | 96 | 0 | 0 |
| ↔ | T2.3_1_M | 1051013092 | 04 / 04 / 2026 | FGM6_70_M | 1051013092 | 04 / 04 / 2026 | No | Yes | 96 | 0 | 0 |
| ↔ | T2.3_1_M | 1051013092 | 04 / 04 / 2026 | FGM6_1_M | 1051013092 | 04 / 04 / 2026 | No | Yes | 96 | 0 | 0 |

**Figure 3. SNP fingerprint matching between different samples from the same individuals.** Samples T2.2 and FGM2 were from the same individual and their fingerprints matched 100%. Samples T2.3, FGM1 and FGM6 were from the same individual and their fingerprints matched 100%.

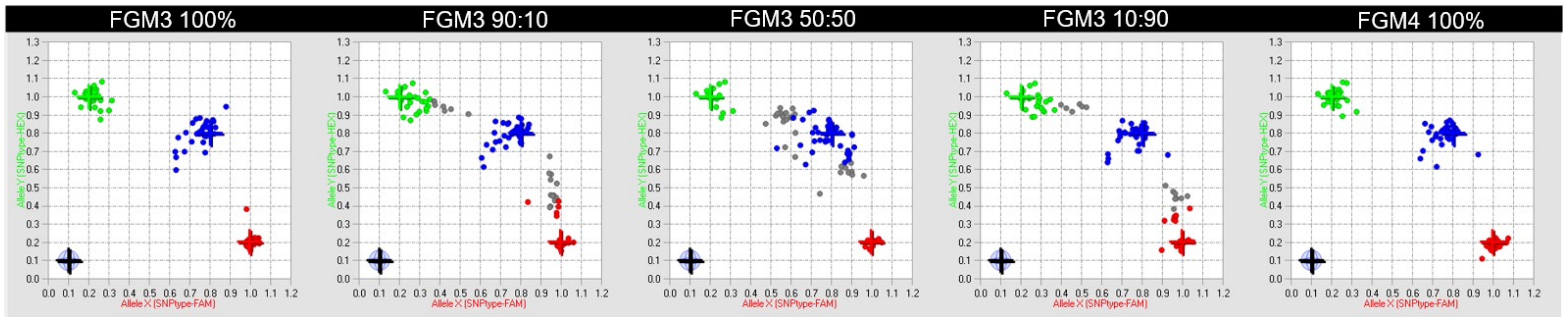


**Figure 2. Cross-contamination samples show increased heterozygous cluster spread in two-dimensional scatter plots.** Two samples (FGM3 and FGM4) were mixed at 90:10, 50:50 and 10:90 ratios by volume and run in duplicate. Only one replicate of each is shown in this figure. As contamination rate increases each cluster spreads out into neighboring clusters, thus decreasing call rate.

## Creating SNP fingerprints to assign individual identity

Genotype calls for each sample and SNP are automatically made with Standard BioTools SNP Genotyping Analysis Software and are presented in a two-dimensional scatter plot for each SNP (Figure 4a). Data is organized into a call map for visualization (Figure 4b). Detailed table data for each run was exported as a .csv file, which was imported into a third-party application to assemble sample-specific SNP fingerprints that were used for comparison against all other samples run within the study.

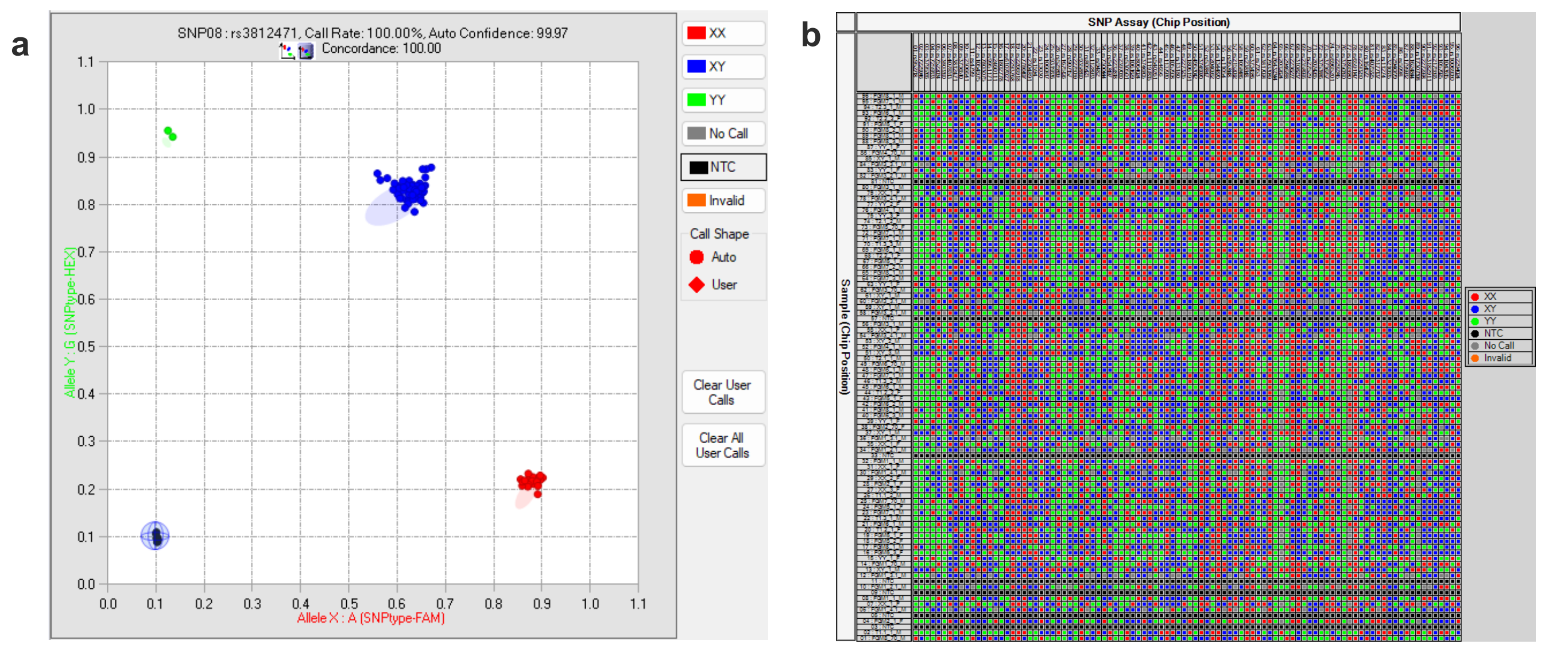


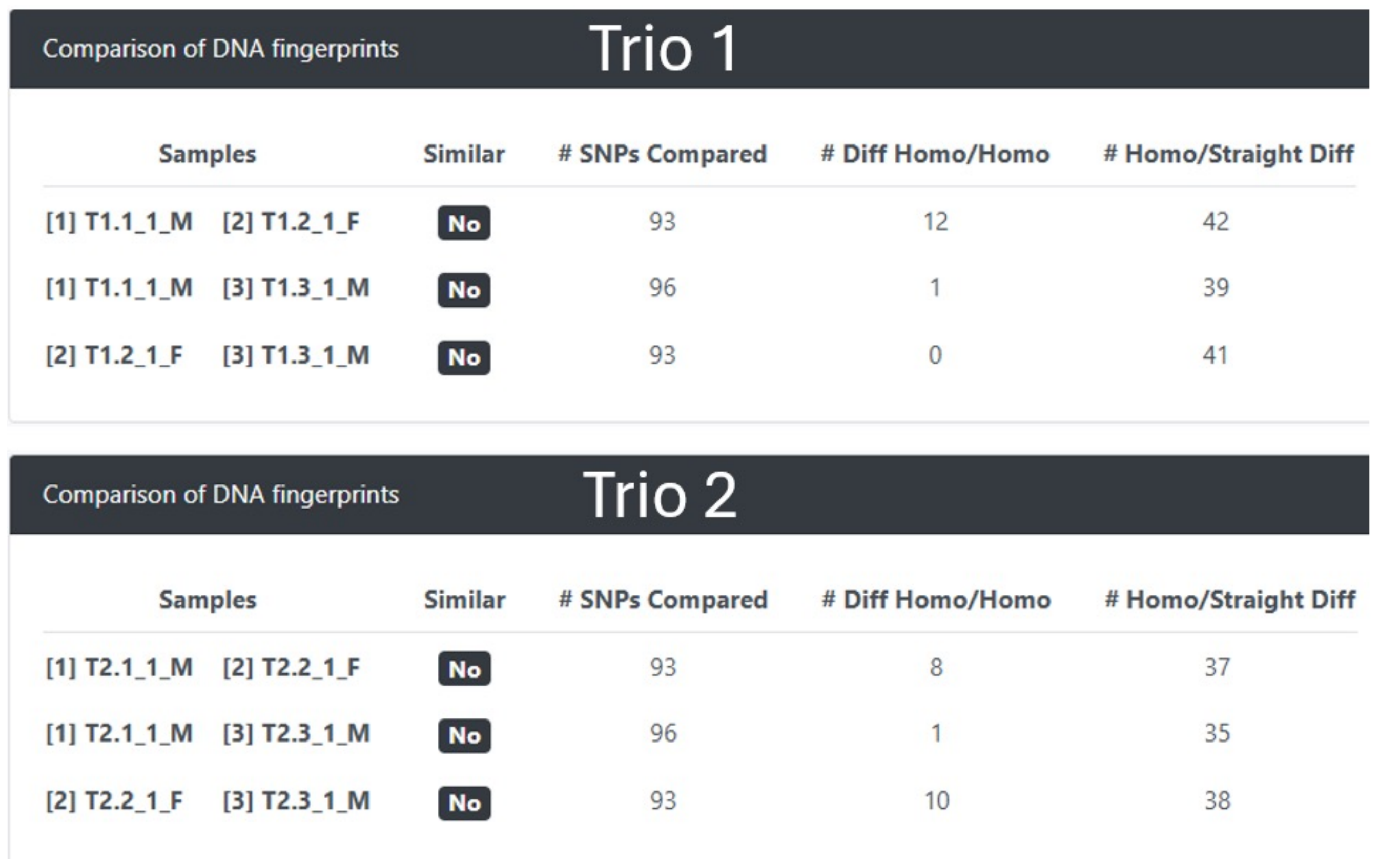
**Figure 4. a) Example plot for one of 96 SNPs used in the Advanta Sample ID Genotyping Panel. b) Example call map for one run totaling 9,216 PCR reactions using the Advanta Sample ID Genotyping Panel.**

## Assessing sample quality

Five human genomic DNA research samples of known quality and concentration were diluted to 10 ng/uL as previously described. An aliquot was removed and stored at 4 °C while the remaining volume of each sample was placed in a thermal cycler set to 95 °C with the heated lid on. Exposure to high heat for an extended period of time should degrade genomic DNA per Bhoyar et al[1]. An aliquot of each sample was removed every 30 minutes up to 120 minutes. Each sample was then run in the IFC in triplicate. Overall call rate remained at 100% for 0–90 minutes and it dropped slightly at 120 minutes (Table 2).

| Sample Names | FGM1-5_0m | FGM1-5_30m | FGM1-5_60m | FGM1-5_90m | FGM1-5_120m |
|---|---|---|---|---|---|
| Average Call Rate % | 100 | 100 | 100 | 100 | 99.9 |

**Table 2. Individual degradation sample call rates used to assess sample quality for SNP genotyping.** Samples were subjected to 95 °C for up to 120 minutes with minimal impact on call rates.



Comparison of DNA fingerprints — Trio 1

| Samples | Similar | # SNPs Compared | # Diff Homo/Homo | # Homo/Straight Diff |
|---|---|---|---|---|
| [1] T1.1_1_M [2] T1.2_1_F | No | 93 | 12 | 42 |
| [1] T1.1_1_M [3] T1.3_1_M | No | 96 | 1 | 39 |
| [2] T1.2_1_F [3] T1.3_1_M | No | 93 | 0 | 41 |

Comparison of DNA fingerprints — Trio 2

| Samples | Similar | # SNPs Compared | # Diff Homo/Homo | # Homo/Straight Diff |
|---|---|---|---|---|
| [1] T2.1_1_M [2] T2.2_1_F | No | 93 | 8 | 37 |
| [1] T2.1_1_M [3] T2.3_1_M | No | 96 | 1 | 35 |
| [2] T2.2_1_F [3] T2.3_1_M | No | 93 | 10 | 38 |

**Figure 5. SNP fingerprint matching between two father-mother-child trios.** None of the samples was deemed genetically identical and could be differentiated from each other.

## Conclusions

Sample degradation, misidentification and contamination are significant risks to the integrity of a biobank. Distribution of compromised samples can waste time and laboratory resources and negatively impact the integrity and reproducibility of research studies.

In this study, the Advanta Sample ID Genotyping Panel was used with the Biomark X9 System to evaluate individual sample quality characteristics by assigning a SNP fingerprint identity, which was used to detect sample cross contamination, assess sample quality and identify samples from the same individuals. This added QC data can inform the biobank staff so critical resources aren't wasted by distributing or further processing compromised samples.

The workflow used in this study demonstrates that the Advanta Sample ID Genotyping Panel paired with the Biomark X9 System and the 96.96 GT Preamp IFC-X offers a fast, efficient and robust workflow that can be deployed as a critical quality-control tool in biorepositories.

---

[1] Bhoyar, L. et al. "An overview of DNA degradation and its implications in forensic caseworks." *Egypt Journal of Forensic Sciences* 14 (2024): 15.

Standard BioTools Inc.
2 Tower Place, Suite 2000, South San Francisco, CA 94080 USA
+1 650 266 6000 • Toll-free in the US and Canada: 866 359 4354
standardbio.com