Human Sample Identification and Mixed Sample Analysis Using the *GenePrint*® 24 System

A Spectrum Compact CE System Application Note

Promega Corporation

Sample Types:
Human gDNA

Instrument Requirements:
- Spectrum Compact CE System (Cat.# CE1304)
- Veriti® Thermal Cycler, 96-Well (Thermo Fisher Scientific Cat.# A48141)

Promega Reagents:
- *GenePrint*® 24 System (Cat.# B1870, B1874)

Other Reagents and Consumables Required:
- Spectrum Compact CE System consumables (cartridge, polymer, septa, retainer, buffer; see Ordering Information at end of document for full list)
- *GenePrint*® 5C Matrix Standard (Cat.# B1930)
- Hi-Di™ Formamide (Thermo Fisher Scientific Cat.# 4401457)

Optional Instruments and Reagents:
- Maxwell® RSC Instrument (Cat.# AS4500)
- Maxwell® RSC Whole Blood DNA Kit (AS1520)
- ProNex® DNA QC Assay (Cat.# NG1004)
- CFX96 Touch Real-Time PCR Detection System (Bio-Rad Cat.# 1855195)

DNA Analysis Software:
- ChimerMarker® Automated Chimerism Analysis Software (SoftGenetics)
- GeneMapper® Software Version 6 (Applied Biosystems) or similar

Introduction

Human sample identification is an essential element of many research projects employing human cells, tissues, or mixtures. Identification may be required to track and confirm sample provenance for databasing, biobanking, or research purposes. For example, archival tumor/normal specimen pairs may be tested to confirm they are derived from the same individual. Human identification may also be used as verification of twin zygosity or parentage in research studies to improve reliability of familial information. In its most sensitive application, human identification may be used to detect mixtures in human research samples—to rule out external contamination in samples, to identify contamination due to histological procedures, to track chimerism, or to study xenografts, among other things.

DNA genotyping using short tandem repeat (STR) markers is currently the gold-standard for human sample identification. STR markers consist of short, repetitive sequence elements 3–7bp in length (1–4). Distributed widely throughout the human genome, STR markers are highly polymorphic genetic markers which may be used for both forensic and non-forensic human identification (5–9). Genotyping at STR markers can be performed using PCR with dye-labeled primers. The number of repeats for each allele in a given individual can be determined, based on the size of the amplified fragments, following electrophoretic separation on a capillary electrophoresis instrument.

Best Practices for Non-Forensic Human Sample Identification

The *GenePrint*® 24 System is a 5-color fluorescent PCR multiplex optimized for human sample identification in research labs. The kit includes 22 highly polymorphic STR markers plus two sex-linked markers. All 24 markers can be amplified in a single PCR reaction, streamlining lab workflows. The high number of markers included in the *GenePrint*® 24 System allows for better discrimination between human samples, including related individuals. Used in detection of human mixtures, this system also increases the likelihood of informative alleles, unique to both contributors, that can be used to estimate their relative contributions to the sample. Many researchers have relied on forensic chemistries that are often optimized for samples with very low DNA concentration. In contrast, the *GenePrint*® 24 System has been optimized for 2.5–5ng DNA input specifically for the research community. The increased DNA input can improve robust detection of minor genetic contributors by reducing stochastic effects on PCR.

The recent launch of the Spectrum Compact CE System now brings the benefits of the *GenePrint*® 24 System and human sample identification to individual labs. The small footprint and low-to medium throughput of this capillary electrophoresis system means that research labs can now perform their own testing without waiting for sample batching or core lab turnaround. Researchers can control the entire workflow, from amplification and electrophoresis conditions, to data analysis, with a potential for faster turnaround to support key research projects. This study aims to demonstrate the utility of the Spectrum Compact CE System paired with the *GenePrint*® 24 System for human identification by focusing on the more sensitive application of human mixture analysis. We demonstrate that the Spectrum Compact CE System can be used for detection of low-level human contributors in mixed samples.

Methods

DNA was purified from whole blood of one female and one male individual using the Maxwell® RSC Whole Blood DNA Kit (Cat.# AS1520) on a Maxwell® RSC Instrument (Cat.# AS4500). Purified DNA was analyzed for purity by absorbance [A260/A280 = 1.94-1.97 and A260/A230 = 1.9-2.0]. DNA was quantified using the qPCR-based ProNex® DNA QC Assay (Cat.# NG1004) on a CFX96 Touch Real-Time PCR Detection System (Bio-Rad).

Male and female DNA were diluted to 10ng/µl in 10mM Tris-HCl, 100µM EDTA, 20µg/ml glycogen (pH 8). Samples were then mixed and diluted further to yield 50%, 20%, 10%, 5%, 3%, 2%, and 1% male DNA in a background of female DNA at 2ng/µl total for use with 5.0ng DNA input, or at 1ng/µl for use with 2.5ng DNA input. Single source DNA and all mixtures were amplified with the *GenePrint*® 24 System (Cat.# B1870) in triplicate with either 2.5ng input and 27 PCR cycles or 5ng input and 26 cycles on a Veriti® Thermal Cycler, 96-Well (Thermo Fisher). Triplicate no template control (NTC) reactions were also included. Amplification products and internal lane standard (WEN ILS 500) were then denatured with Hi-Di™ formamide (Thermo Fisher), including at least one allelic ladder per 16 samples (4 injections). Samples were electrophoresed on the Spectrum Compact CE System with Polymer7 using the pre-loaded assay settings for Promega_5Dye_WENILS_36_P7. Injection and run conditions are given in Table 1. All NTC reactions were blank.

For qualitative analysis and peak height ratio analysis, data were analyzed in GeneMapper™ Software v6. Artifacts, bleed through, and n+1 repeat stutter (e.g. n+3nt, n+4nt, or n+5nt) above the analytical threshold (75RFU) were omitted manually. For these studies a 75RFU threshold was determined to be suitable for the *GenePrint*® 24 System on the Spectrum Compact CE System. This was determined using a serial dilution of male and female DNA amplified in triplicate and injected with the same parameters used here (data not shown). All data was analyzed using GeneMapper™ Software. Peak height ratio was calculated for male and female single source samples as the smaller peak height divided by the taller peak height x 100% for heterozygous markers. Homozygous markers were noted as 100% by default. Data were analyzed for the percent of unshared minor contributor alleles detected (i.e. number of male-specific alleles > 75RFU), regardless of zygosity. All data were reported as the mean ± standard deviation across triplicate reactions.

Semi-quantitative analysis of the mixture percentages was performed in ChimerMarker™ software v3.1.5 (SoftGenetics) using Type 1 and Type 2 markers only, including stutter adjustment for markers with informative alleles in stutter positions, and omitting marker D19S433 and the sex-linked markers AMEL and DYS391 (10–12). Please refer to the ChimerMarker™ software User Manual for details on calculations performed by the software.

Results

The Spectrum Compact CE System is a new capillary electrophoresis instrument, featuring a small footprint and easy-to-install consumables, that is capable of both fragment analysis and sequencing applications. This four capillary low-to-medium throughput instrument is ideally suited for human mixed sample analysis applications requiring fast turnaround with minimal batching. We demonstrate here the utility of the Spectrum Compact CE System for sensitive detection of mixed samples with the *GenePrint*® 24 System.

The *GenePrint*® 24 System is a multiplex fluorescent PCR system, allowing single tube amplification and five-color detection of 22 polymorphic short tandem repeat (STR) markers and two sex markers, DYS391 and Amelogenin. The *GenePrint*® 24 System is optimized for 2.5-5.0ng input to improve detection of minor contributors by reducing stochastic effects relative to forensic STR chemistries. Amplified STR alleles are detected by capillary electrophoresis using a polymer with high resolving power, like Polymer7.

Human mixtures were mimicked by combining purified DNA from one male and one female individual at different ratios

down to 1% male : 99% female. Single source DNA and all mixtures were amplified in triplicate with the *GenePrint*® 24 System, as indicated in the technical manual, using either 2.5ng or 5.0ng of input DNA and 27 or 26 PCR cycles, respectively. Amplified samples were denatured, and injected with a sizing ladder on the Spectrum Compact using the default capillary electrophoresis conditions (Table 1). At least one allelic ladder was included per 4 injections (16 samples) for appropriate allele determination.

Table 1. Default fragment analysis parameters for Promega 5-dye chemistries with Polymer 7 on the Spectrum Compact CE System. These run parameters are preloaded as the assay Promega_5Dye_WENILS_36_P7.

Injection Voltage	1.6 kV
Injection Time	9 s
Run Voltage	13 kV
Run Time	1290 s
Oven Temperature	60°C

Data were analyzed in two ways—using a qualitative approach typical of cell line contamination testing or identification of extraneous tissues in tissue samples, and using a quantitative analysis commonly used in chimerism determination.

For qualitative testing, data were analyzed in GeneMapper™ Software v6 and alleles were determined relative to an allelic ladder (Fig.1). A representative portion of the electropherograms is shown in Figure 2. Marker D1S1656 demonstrates how relative peak heights change with allele zygosity and mixture ratios. In this case, the female is heterozygous with peak height evenly distributed between two alleles (Fig.2, panel 1). The male genotype is homozygous with a single peak of approximately 2-fold higher fluorescence (Fig.2, panel 4). When male and female DNA is mixed evenly (Fig.2, panel 2), all peak heights are reduced with ratios of approximately 1:1:2. As the percent of the male DNA decreases (Fig.2, panel 3), the peak height of the male-specific allele also decreases.

Sample mixtures were qualitatively identified by detection of alleles unique to the minor contributor (i.e. male). Genotypes for both the male and female individuals are given in Table 2, with unique male alleles highlighted in red. For this pair of samples, the male DNA introduces 30 unique alleles across the 24 loci amplified with the *GenePrint*® 24 System. Half of these unique alleles can be detected above the analytical threshold with as little as 3% male DNA in the mixture, demonstrating sensitive detection of a minor contributor (Fig.3).

For semi-quantitative testing, the same data were analyzed using ChimerMarker™ software in a workflow mimicking detection of chimerism. Calculations of chimerism are based on relative peak heights of the two samples in the mixture, assuming that peak height is proportional to the percent

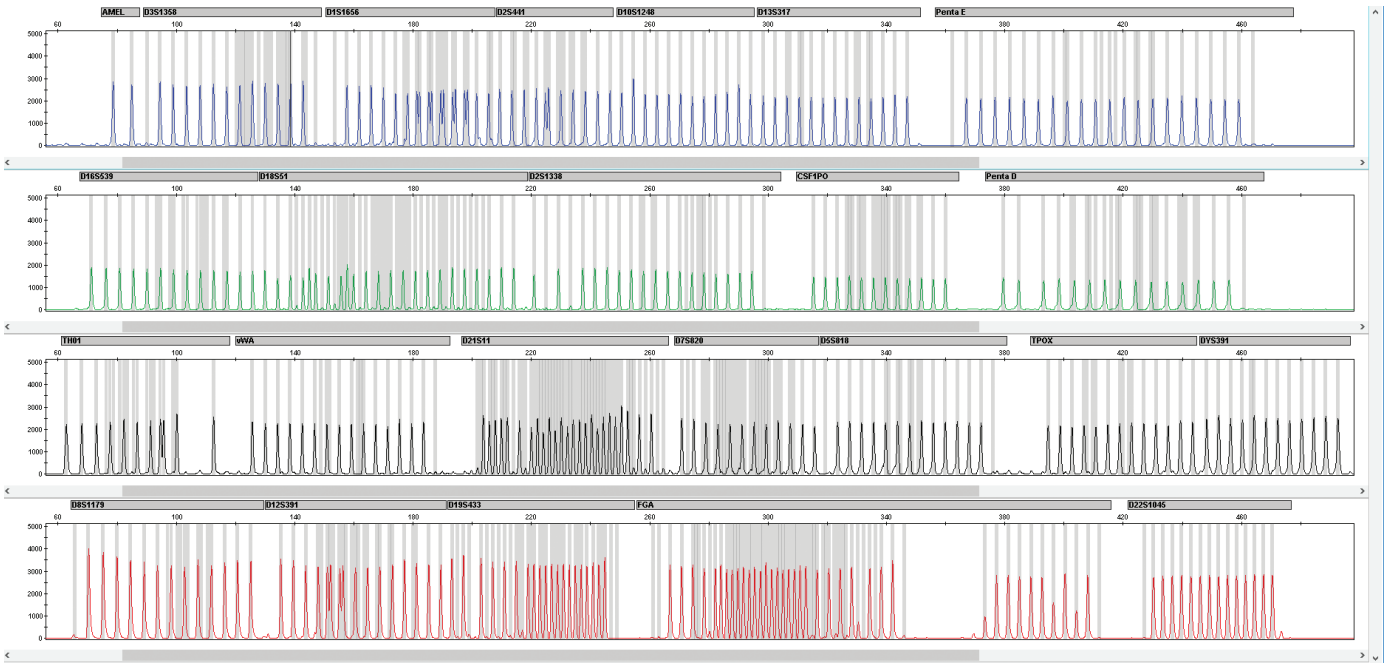

Figure 1. The *GenePrint*® 24 System allelic ladder analyzed on a Spectrum Compact CE System.

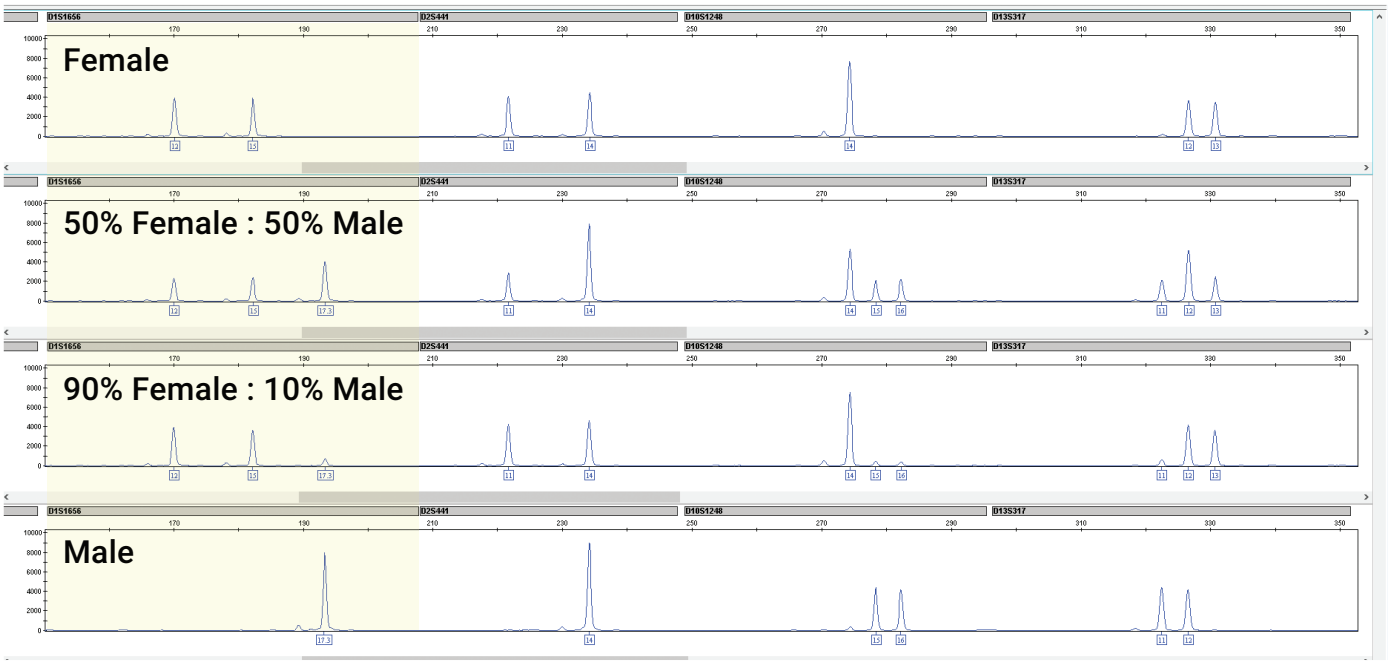

Figure 2. Representative FAM (blue) channel electropherograms showing unique male alleles detected in a mixture with a female sample, amplified with the *GenePrint*® 24 System and analyzed on the Spectrum Compact CE System. Highlighted in yellow is the marker D1S1656.

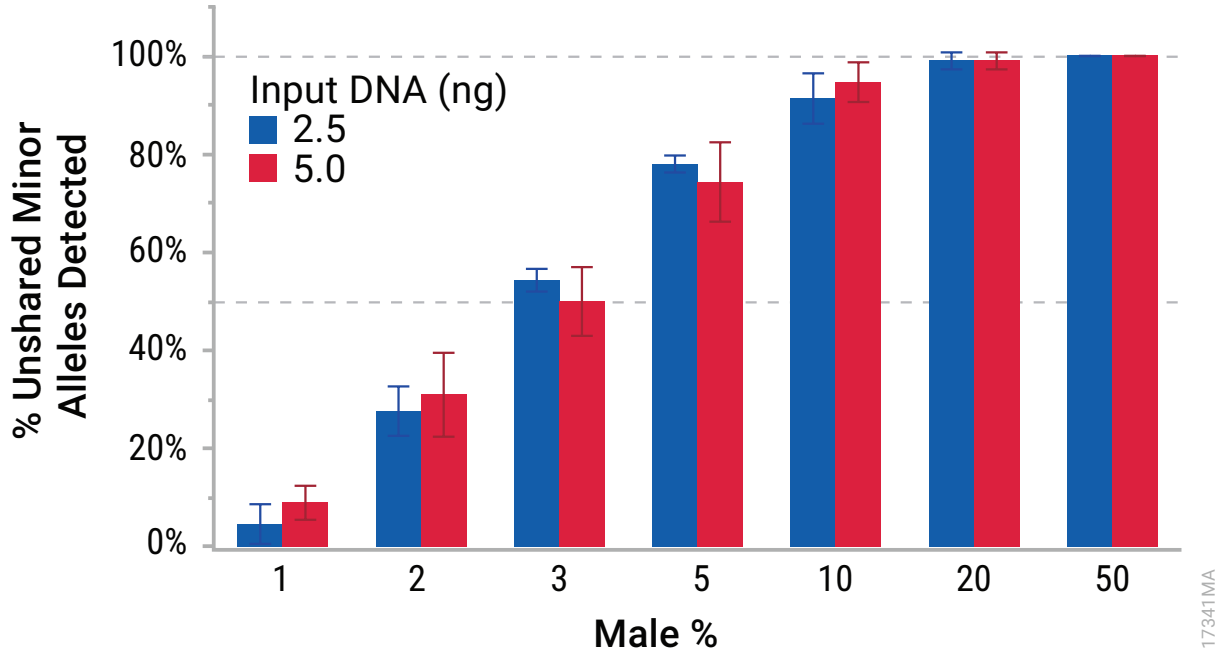

Figure 3. Detection of unshared minor alleles in mixed samples amplified with the *GenePrint*® 24 System and analyzed on a Spectrum Compact CE System. All mixtures of male and female DNA were amplified in triplicate using either 2.5ng input or 5.0ng input. The percent of unique male alleles detected (out of 30 possible, relative to the female sample) is shown as the mean ± standard deviation of n=3.

of each allele in the PCR reaction (11). To verify that the *GenePrint*® 24 System and the Spectrum Compact CE System generate proportional peaks heights, peak height ratios were calculated for the male and female single source samples. The average peak height ratio was >80% for all markers for both male and female samples, as shown in Figure 4.

Markers were characterized for use in chimerism quantification according to guidelines issued by the

Table 2. Genotypes of female and male DNA used for mixed sample analysis and characterization of markers for quantitative analysis of chimerism. Markers TH01 and D8S1179 were excluded from chimerism calculations due to inconsistent baseline subtraction using the ChimerMarker™ software. Markers D22S1045 and D19S433 were excluded due to highly variable and/or outlier estimates of chimerism in the data set.

Marker	Female genotype	Male genotype	Marker type	Informative allele in stutter position	Used to calculate % chimerism
AMEL	X	X,Y	X/Y		No
CSF1PO	11	11,12	3		No
D1S1656	12,15	17.3	1	No	Yes
D2S441	11,14	14	Uninformative		No
D2S1338	18	21,24	1	No	Yes
D3S1358	15,17	14,17	2	Yes	Yes
D5S818	12,13	10,11	1	Yes	Yes
D7S820	9,10	8,11	1	Yes	Yes
D8S1179	13,15	10	1	No	No
D10S1248	14	15,16	1	Yes	Yes
D12S391	17.3,21	16,20.3	1	No	Yes
D13S317	12,13	11,12	2	Yes	Yes
D16S539	11	13	1	No	Yes
D18S51	10,13	16,19	1	No	Yes
D19S433	12,14	13,14	2	Yes	No
D21S11	30,31	29	1	Yes	Yes
D22S1045	16	14,15	1	Yes	Yes
DYS391	N/A	10	X/Y		No
FGA	21,22	21,22	Uninformative		No
Penta D	9,13	9,13	Uninformative		No
Penta E	7,12	11,12	2	Yes	Yes
TH01	6,7	9,10	1	No	No
TPOX	11	8	1	No	Yes
vWA	18,19	14,16	1	No	Yes

United Kingdom National External Quality Assessment Service (UK NEQAS, 11–12). Markers D1S1656 and D10S1248 are examples of fully informative Type 1 markers, with no shared alleles between the female and male genotypes (Table 2, Figure 2). Marker D13S317 is an example of an informative Type 2 marker, with heterozygous genotypes for both female and male samples and one shared allele. Marker D2S441 is an example of a Type 3 marker with at least one individual—in this case the male—having no unique alleles. Full genotypes are given for the female and male samples in Table 2, along with characterization of each marker type for chimerism quantification.

Consistent with best practices recommended by the UK NEQAS guidelines, only Type 1 and Type 2 informative markers with unique alleles for both samples were used to calculate chimerism (11). Marker D19S433 was excluded due to inability to discern female allele stutter from the male allele. For markers with informative male alleles in stutter positions relative to female alleles, the stutter adjusted % chimerism values provided by ChimerMarker™ software were used (10).

Percent chimerism was plotted relative to known percent of the female DNA in the sample mixture (Fig.5). Agreement is linear between the experimentally measured percent chimerism and the known mixture percent. The *GenePrint*® 24 System can be paired with the Spectrum Compact CE System for sensitive qualitative or quantitative detection of human mixed samples.

Conclusion

Human sample identification and mixed sample analysis with STR genotyping can confirm sample provenance and identify sample contamination or chimerism in many research areas. Now the Spectrum Compact CE System can be paired with the *GenePrint*® 24 System to allow individual labs, or small lab clusters, control over their sample testing. The low-to-medium throughput design of the Spectrum Compact CE System and accompanying reagents mean samples can be run quickly–without batching samples and without reagent waste. Together with the *GenePrint*® 24 System, the Spectrum Compact CE System can enable detection of a low percentage minor contributor in a human mixture, demonstrating utility for even the most sensitive applications.

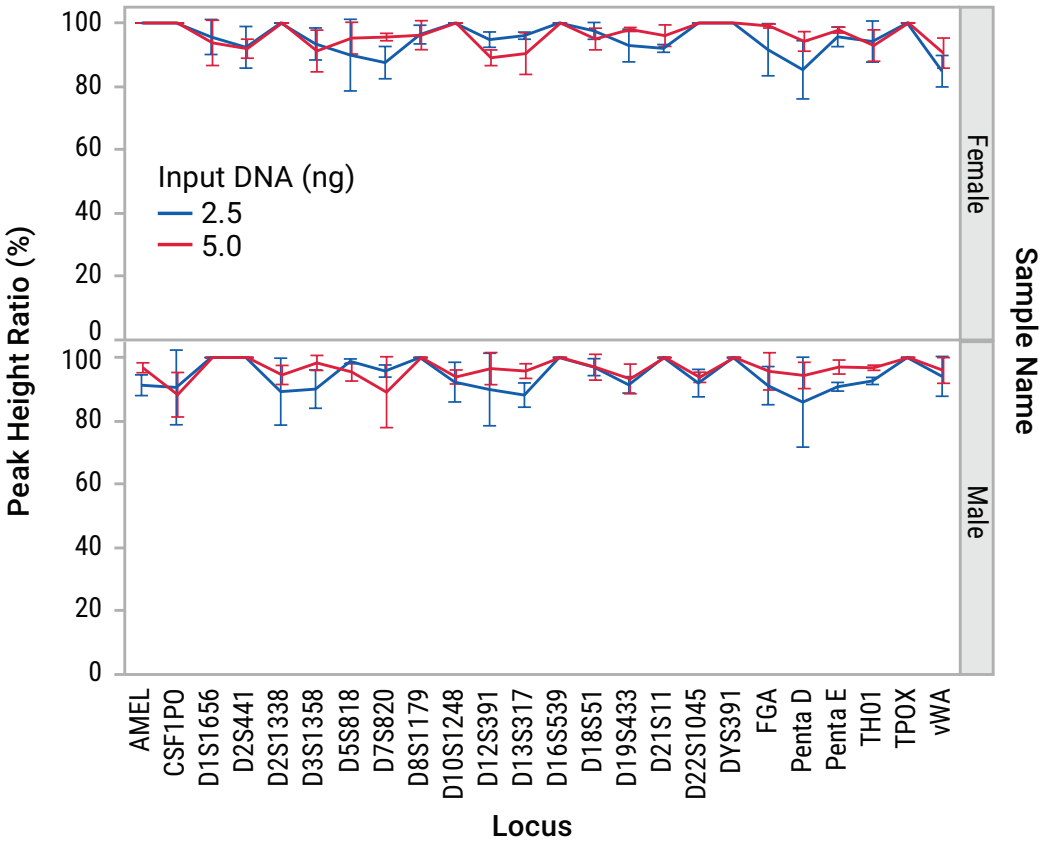

Figure 4. Peak height ratios for female (top) and male (bottom) single source DNA samples amplified with the *GenePrint*® 24 System and injected on a Spectrum Compact CE System. Peak height ratios for all heterozygous loci were calculated as the shorter peak height divided by the taller peak height and expressed as a percent. Mean ± standard deviation of n=3, shown. Homozygous markers are shown as 100% by default. The Y chromosome marker DYS391 was omitted.

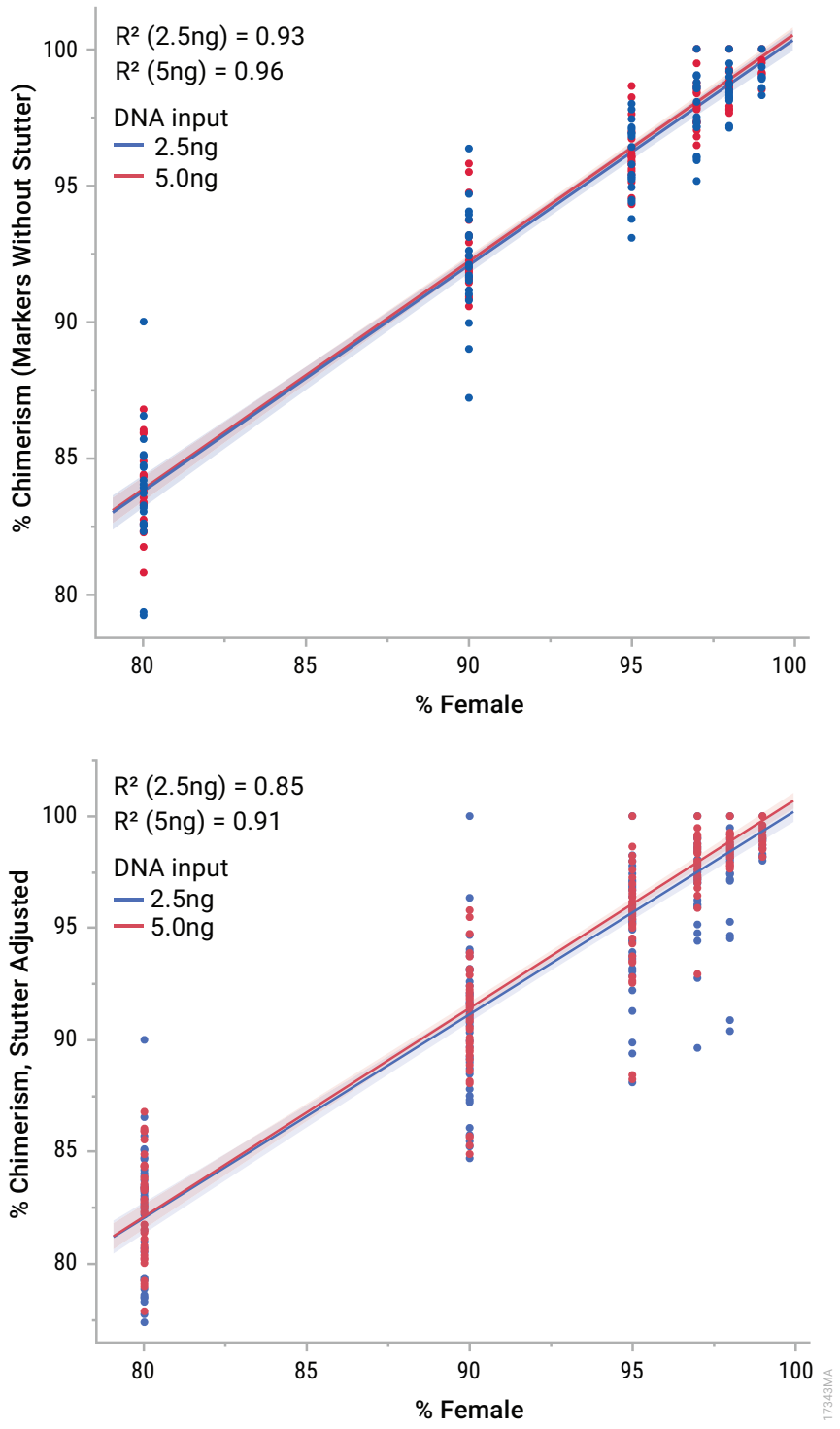

Figure 5. Percent chimerism calculated for mixtures of male and female DNA, amplified with the *GenePrint*® 24 System and injected on a Spectrum Compact CE System. (Top) Percent chimerism at each marker was calculated using ChimerMarker™ software for the 7 informative Type 1 and Type 2 markers not impacted by stutter—D1S1656, D16S539, D18S51, D2S1338, vWA, TPOX, and D12S391. (Bottom) Stutter adjusted percent chimerism was calculated using ChimerMarker™ software for the 14 informative Type 1 and Type 2 markers (excepting TH01, D8S1179, D22S1045, and D19S433). Percent chimerism is shown for each marker and replicate (n=3) and a linear regression with 95% confidence intervals was fitted to the data. R^2 values are reported.

References

1. Edwards, A. *et al.* (1991) DNA typing with trimeric and tetrameric tandem repeats: Polymorphic loci, detection systems, and population genetics. In: *The Second International Symposium on Human Identification 1991*, Promega Corporation, 31–52.
2. Edwards, A. *et al.* (1991) DNA typing and genetic mapping with trimeric and tetrameric tandem repeats. *Am. J. Hum. Genet.* **49**, 746–56.
3. Edwards, A. *et al.* (1992) Genetic variation at five trimeric and tetrameric tandem repeat loci in four human population groups. *Genomics* **12**, 241–53.
4. Warne, D. *et al.* (1991) Tetranucleotide repeat polymorphism at the human b-actin related pseudogene 2 (actbp2) detected using the polymerase chain reaction. *Nucleic Acids Res.* **19**, 6980.
5. Ausubel, F.M. *et al.* (1996) Unit 15: The polymerase chain reaction. In: *Current Protocols in Molecular Biology*, Vol. 2, John Wiley and Sons, NY.
6. Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) Chapter 14: In vitro amplification of DNA by the polymerase chain reaction. In: *Molecular Cloning: A Laboratory Manual*, Second Edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.
7. *PCR Technology: Principles and Applications for DNA Amplification* (1989) Erlich, H.A., ed., Stockton Press, New York, NY.
8. *PCR Protocols: A Guide to Methods and Applications* (1990) Innis, M.A. *et al.* eds., Academic Press, San Diego, CA.
9. Butler, J.M. (2005) *Forensic DNA Typing*, 2nd ed., Elsevier Academic Press, London.
10. ChimerMarker™ Automated Chimerism Analysis Software User Manual (2016) SoftGenetics, LLC, State College, PA.
11. Clark, J.R., *et al.* (2014) Monitoring of chimerism following allogeneic haematopoietic stem cell transplantation (HSCT): technical recommendations for the use of short tandem repeat (STR) based techniques, on behalf of the United Kingdom National External Quality Assessment Service for Leucocyte Immunophenotyping Chimerism Working Group. *British Journal of Haematology* **168**, 26–37.
12. Thiede, C., *et al.* (1999) Rapid quantification of mixed chimerism using multiplex amplification of short tandem repeat markers and fluorescence detection. *Bone Marrow Transplantation* **23**, 1055–60.

Ordering Information

Product	Cat.#
Spectrum Compact CE System*	CE1304
Spectrum Compact Capillary Cartridge, 4-Capillary, 36cm*	CE2340
Spectrum Compact Polymer4*	CE2304
Spectrum Compact Polymer7*	CE2307
Spectrum Compact Buffer*	CE2300
Spectrum Compact Cathode Septa Mat*	CE2301
Spectrum Compact Cathode Retainer*	CE2302
Spectrum Compact Strip Base & Retainer, 32-Well*	CE2332
Strip Septa Mat, 8-Well*	CE2308
GenePrint® 24 System*	B1870, B1874
GenePrint® 5C Matrix Standard**	B1930
Maxwell® RSC Instrument	AS4500
Maxwell® RSC Whole Blood DNA Kit	AS1520

*Not for Medical Diagnostic Use.

**For In Vitro Research Use Only. Not for Diagnostic Procedures.

PROMEGA CORPORATION • 2800 WOODS HOLLOW ROAD • MADISON, WI 53711-5399 USA • TELEPHONE 608-274-4330
www.promega.com • © 2020 PROMEGA CORPORATION • ALL RIGHTS RESERVED • PRICES AND SPECIFICATIONS SUBJECT TO CHANGE WITHOUT PRIOR NOTICE • PRINTED IN USA 11/20 • 59894754 • PART #AN361