Systematic Analysis of Interference Due to Stutter in Estimating Chimerism Following Hematopoietic Cell Transplantation

Bharat Thyagarajan,^{1*} Shawn Young,² Stacy Floodman,²
Randolph Peterson,¹ and Xinjing Wang¹

¹Department of Laboratory Medicine and Pathology, University of Minnesota, Minneapolis, Minnesota

²University of Minnesota Medical Center, Fairview, Molecular Diagnostics Laboratory, Minneapolis, Minnesota

Stutter formed during amplification of short tandem repeats (STRs) interfere with accurate engraftment monitoring following hematopoietic cell transplantation (HCT). We describe a mathematical approach to minimize the contribution of stutter when estimating chimerism following HCT. Pretransplant DNA samples from 409 donors and recipients were used to define marker-specific stutter cut-off values for all makers used in the AMPF/STR Profiler Plus amplification kit. Mock chimerism samples (5, 20, and 50%) were used to evaluate the contribution of stutter in estimating chimerism. Three markers, vWA, D13S317, and D18S51, had overlapping stutter in the mock chimerism

samples. Only D18S51 had stutter from a shared allele, whereas the other two markers had stutter from nonshared alleles. Without adjustment for stutter, D18S51 showed a 8, 6, and 4% difference from expected chimerism values (5, 20, and 50%) and after adjusting for laboratory-defined marker-specific stutter cut-offs the corresponding difference from expected chimerism values was 1, 0, and 0%. Difference from expected chimerism values in the vWA and D13S317 markers were similar before and after adjustment of stutter. Adjustment for stutter from shared alleles may improve accuracy of estimated chimerism. *J. Clin. Lab. Anal.* 23:308–313, 2009. © 2009 Wiley-Liss, Inc.

Key words: stutter; chimerism; hematopoietic cell transplantation; STR markers; accuracy

INTRODUCTION

Monitoring engraftment following hematopoietic cell transplantation (HCT) is necessary for early documentation of engraftment, detection and timely treatment of graft rejection or early relapse (1). Polymorphic markers, such as variable nucleotide tandem repeats (VNTRs), are commonly used to determine chimerism following HCT (2). There are two major types of VNTRs: minisatellites (nucleotide repeats ranging from 6 to 50 nucleotides) and microsatellites/short tandem repeats (STRs) (nucleotide repeats ranging from 2 to 5 nucleotides) (3). Commercial STR kits that amplify multiple STR markers in a single reaction and allow small amounts of DNA to be used for identity testing have greatly facilitated the use of STR makers in clinical laboratories to estimate chimerism after HCT. In addition, STRs have shown to be almost equally informative as minisatellites in being able to distinguish

between donors and recipient (4). However, a major limitation of all STR assays is the presence of stutter, which is typically one repeat length smaller than the parent nominal allele and is generated by slippage of the DNA during polymerase chain reaction (PCR) (5,6). Co-localization of stutter from nominal allele of one individual with nominal allele from another individual potentially limits the usefulness of STR markers in uniquely identifying individuals (7). A common strategy to avoid this problem is to choose STR markers that uniquely identify donors and recipient without interference from stutter (2). In cases where there are

*Correspondence to: Bharat Thyagarajan, Department of Laboratory Medicine and Pathology, 1-136 Moos Towers, 420, Delaware Street SE, Minneapolis, Minnesota 55455. E-mail: thya0003@umn.edu

Received 15 December 2008; Accepted 15 July 2009

DOI 10.1002/jcla.20336

Published online in Wiley InterScience (www.interscience.wiley.com).

multiple donors in a recipient sample (e.g., double umbilical cord blood transplants or multiple HCT), excluding markers owing to the presence of stutter may result in inability to quantitate all components (recipient and donors) using a single marker or a uniform set of markers. As experimental strategies have proven to be of limited usefulness in eliminating stutter (6), mathematical approaches to reduce the contribution of stutter peaks in determining chimerism can provide useful information in this regard. One recent article demonstrated the usefulness of a mathematical approach to minimize the contribution of stutter in chimerism analyses (8).

In this article we describe and validate a mathematical approach to minimize the effect of stutter when estimating chimerism in a clinical molecular diagnostics laboratory.

MATERIAL AND METHODS

Biological Samples

Pretransplant DNA samples from 409 consecutive donors and recipients were analyzed to determine the range of stutter for nine markers. Twenty post-HCT patient samples with percentage recipient ranging from 0 to 96% were used to evaluate accuracy and reproducibility of the proposed data analysis algorithm to analyze clinical samples that have a range of chimerism values.

Preparation of mock chimerism samples to determine sensitivity and reliability of chimerism analyses

DNA obtained from two individuals who were polymorphic for all markers except amelogenin was

mixed in different proportions, after standardizing the DNA concentrations, to yield 5, 20, and 50% chimerism. These samples were used to estimate chimerism in nine markers and evaluate the sensitivity, accuracy, and reliability of the modified data analysis procedures to estimate chimerism.

STR Kit

The AmpF ℓ STR Profiler Plus amplification kit (Applied Biosystems Inc., Foster City, CA) containing nine polymorphic STR markers (Table 1) and the amelogenin marker was used to evaluate engraftment in samples from patients who underwent HCT. All samples were amplified in a 50 μ L reaction containing 19.1 μ L of the AmpF ℓ STR Reaction Mix, 10.0 μ L of AmpF ℓ STR primer set, 0.9 μ L of the Amplitaq Gold Polymerase, and 20.0 μ L of DNA sample (DNA concentration = 0.1 ng/ μ L). The cycling conditions were initial denaturation at 95°C for 11 min followed by amplification step that consisted of 94°C for 1 min, 59°C for 1 min, and 72°C for 1 min for 28 cycles. This was followed by a final elongation step at 60°C for 45 min. Capillary electrophoresis was performed using a 3130xl DNA Sequencing System (Applied Biosystems Inc.) and data was analyzed using Genemapper v4.0 software (Applied Biosystems Inc.).

STATISTICAL ANALYSIS

Descriptive statistics were obtained using MS EXCEL 2007 (Microsoft Corp, Seattle, WA). Weakly amplified reactions, defined as reactions having nominal peak

TABLE 1. Distribution of Markers in the Profiler Plus Kit and Comparison of Laboratory-Defined and Manufacturer-Defined Cut-Off Values

STR locus	Fluorescent dye	Repeat structure (number of repeats)	Median (range) (%)	Laboratory-defined stutter cut-off limit (%)	Manufacturer-defined stutter cut-off limit (%)
D3S1358	FAM	TCTA-TCTG (12-19)	6.6 (3.3-16.3)	7	11
vWA	FAM	TCTA-TCTG (11-21)	6.7 (1.3-14.7)	7	11
FGA	FAM	TTTC-TTTT (18-30)	6.6 (1.7-19.8)	7	11
D8S1179	JOE	TCTA (8-19)	5.7 (1.4-14.3)	6	12
D21S11	JOE	TCTA (24.2-38)	6.3 (3.7-14.9)	6	13
D18S51	JOE	AGAA (9-26)	7.6 (3.6-15.1)	8	16
D5S818	NED	AGAT (7-16)	4.9 (1.8-15.9)	5	10
D13S317	NED	TATC (8-15)	4.9 (1.5-8.6)	5	10
D7S820	NED	GATA (6-15)	4.3 (1.9-7.4)	4	9

areas below 2550 relative fluorescence units, were excluded from the stutter analyses. Stutter was defined as the presence of an amplification peak one repeat size lower than the nominal peak(s). The percentage stutter was calculated using the formula (Stutter peak area/Nominal peak area)*100. As the stutter was not normally distributed median and range were calculated for all markers. The median stutter in each marker was used to determine the marker-specific stutter cut-off value.

The mock chimerism samples (5, 20, and 50%) were analyzed across 44–50 consecutive runs to determine the accuracy and reliability of all markers to detect different degrees of chimerism. Means and standard deviations of chimerism across all markers were evaluated in the mock chimerism samples and these samples were also used to evaluate the accuracy and reliability of mathematical correction for stutter peaks. Predetermined marker-specific stutter cut-off value (S) was subtracted from the peak area of the nominal allele with overlapping stutter for adjustment of stutter (Fig. 1). The cut-off values were determined using two methods: (1) Laboratory-defined cut-off values: based on marker-specific median stutter values obtained from the Molecular Diagnostics Laboratory at the University of Minnesota Medical Center, Fairview and (2) Manufacturer-defined cut-off values. Finally, analyses with and without adjustment for stutter

were performed on patient samples to determine the contribution of stutter to error in determining donor chimerism in post-HCT samples.

RESULTS

Stutter for Markers in the AMPF[®]STR Profiler Plus Kit

The peak areas of the nominal alleles were used to estimate the percentage stutter for all nine markers except amelogenin in 409 patient samples. Percentage stutter varied significantly between markers (Table 1). The markers on the NED dye (D5S818, D13S317, and D7S820) had lower percentage stutter (Median: 4.3–4.9% across three markers) as compared with markers on the FAM or JOE dyes (Median: 5.7–7.6% across six markers). Notably FGA, labeled with the FAM dye had wider distribution of percentage stutter as compared with the other markers (Median (range): 6.6% (1.7–19.8%)) (Table 1).

Determination of Chimerism in Mock Chimerism Samples

The average chimerism for all nine markers, unadjusted for contributions from overlapping stutter, across

Fig. 1. Diagrammatic representation of markers with stutter from shared alleles and nonshared alleles.

the three mock chimerism samples is shown in Table 2. Three markers, vWA, D13S317, and D18S51, had overlapping stutter. Of these, D13S317 and vWA had stutter from nonshared alleles, whereas D18S51 had stutter from shared alleles (Fig. 1). Without adjustment for stutter, vWA showed a 3, 0, and -3% difference, D13S317 showed a 2, 1, and 0% difference and D18S51 showed a 8, 6, and 4% difference from expected chimerism values (5, 20, and 50%) (Table 2).

Use of Marker-Specific Stutter Cut-Off Values to Estimate Chimerism

After adjustment for stutter using laboratory-defined cut-offs, vWA showed a 0, -2, and -3% difference, D13S317 showed a 0, -1, and -1% difference and D18S51 showed a 1, 0, and 0% difference from expected chimerism values (5, 20, and 50%) (Table 3). After similar adjustment using manufacturer-defined cutoffs, vWA showed a -2, -3, and -3% difference, D13S317 showed a -5, -4, and -3% difference and D18S51 showed a -2, -3, and -2% difference from expected

chimerism values (5, 20, and 50%) (Table 3). Comparison of these values with estimated chimerism before adjustment for stutter showed that difference from expected values was substantially lower after adjustment for stutter for D18S51 but not for the vWA or D13S317 markers. Use of manufacturer-defined cut-off values resulted in inability to detect 5% chimerism in 24% (12/49) of analytical runs for vWA and in 100% (46/46) runs for D13S317. Use of laboratory-defined cut-offs resulted in detection of 5% chimerism in all analytical runs for vWA, D13S317, and D18S51.

Contribution of Stutter in Estimating Chimerism in Patient Samples

Using % recipient obtained from markers without overlapping stutter as the reference, markers with overlapping stutter from a shared allele between recipient and donor displayed variable amount of bias in estimating % recipient (-6 to 14%) (Fig. 2). Using laboratory- or manufacturer-defined stutter cut-off values to adjust for stutter resulted in reduction in bias

TABLE 2. Observed and Expected Chimerism Values for All Markers (Except Amelogenin) in the Profiler Plus Kit at Three Chimerism Levels (5, 20, and 50%)

STR marker	Expected chimerism (%)	Observed chimerism (Mean $\pm$ 2 SD) (%)	Expected chimerism (%)	Observed chimerism (Mean $\pm$ 2 SD) (%)	Expected chimerism (%)	Observed chimerism (Mean $\pm$ 2 SD) (%)
D3s1358	5	5 $\pm$ 2	20	19 $\pm$ 3	50	48 $\pm$ 4
vWA ^a	5	8 $\pm$ 2	20	20 $\pm$ 3	50	48 $\pm$ 4
FGA	5	6 $\pm$ 3	20	20 $\pm$ 4	50	50 $\pm$ 6
D8s1179	5	5 $\pm$ 3	20	19 $\pm$ 5	50	48 $\pm$ 6
D21s11	5	6 $\pm$ 3	20	19 $\pm$ 5	50	49 $\pm$ 6
D18S51 ^a	5	13 $\pm$ 3	20	26 $\pm$ 4	50	54 $\pm$ 6
D5s818	5	6 $\pm$ 2	20	21 $\pm$ 4	50	50 $\pm$ 5
D13S317 ^a	5	7 $\pm$ 2	20	21 $\pm$ 3	50	50 $\pm$ 4
D7s820	5	6 $\pm$ 2	20	21 $\pm$ 5	50	51 $\pm$ 6

^aThese markers had overlapping stutter peaks when determining chimerism. All chimerism values were determined without any adjustment for stutter.

TABLE 3. Adjustment for Stutter Using Different Cut-Off Values and Comparison With Expected Chimerism Values

STR marker	Expected chimerism (%)	Observed chimerism with no stutter adjustment (mean $\pm$ 2 SD) (%)	Observed chimerism with lab-defined cutoff (mean $\pm$ 2 SD) (%)	Observed chimerism with manufacturer cutoff (mean $\pm$ 2 SD) (%)
vWA	5	8 $\pm$ 2	5 $\pm$ 2	3 $\pm$ 4
vWA	20	20 $\pm$ 4	18 $\pm$ 3	17 $\pm$ 3
vWA	50	47 $\pm$ 2	47 $\pm$ 5	47 $\pm$ 5
D18s317	5	13 $\pm$ 2	6 $\pm$ 3	3 $\pm$ 3
D18s317	20	26 $\pm$ 4	20 $\pm$ 4	17 $\pm$ 3
D18s317	50	54 $\pm$ 6	50 $\pm$ 6	48 $\pm$ 6
D13s51	5	7 $\pm$ 2	5 $\pm$ 2	0 $\pm$ 0
D13s51	20	21 $\pm$ 4	19 $\pm$ 3	16 $\pm$ 4
D13s51	50	50 $\pm$ 4	49 $\pm$ 4	47 $\pm$ 4

Fig. 2. Bias in estimating chimerism using markers that have stutter from shared alleles in 20 patient samples obtained after hematopoietic cell transplantation (HCT). All markers except D18S51 and D13S317 had stutter from shared alleles in the 20 samples.

in a majority of samples (16/20 (80%) and 14/20 (70%), respectively) (Fig. 2).

DISCUSSION

This study confirms that mathematical adjustment for stutter using marker-specific stutter cut-off values can improve accuracy of estimated chimerism.

An important finding from this study is that only stutter from shared alleles contributes substantially to error in estimating chimerism, whereas stutter from nonshared alleles did not substantially affect accuracy of estimated chimerism. This finding is consistent with results from a previous study by Chen et al., which also used mathematical approaches to minimize the contribution of stutter in chimerism analysis (8). In the study by Chen et al. only markers with stutter from shared alleles (D3S1358 and vWA) showed substantial improvement in estimated chimerism values after adjustment for stutter (observed chimerism in a 5% mock chimerism sample was 11.2 vs. 4.1% using D3S1358 before and after adjustment for stutter, respectively, whereas corresponding values for vWA were 10.9 vs. 2.8%) (8). In contrast, the D7S820 marker that had stutter from nonshared alleles showed an increase in bias at the 5% level after adjustment for stutter (observed chimerism in a 5% mock chimerism sample was 5.6% before adjustment for stutter vs. 2.8% after adjustment for stutter) (8). Thus, results from this and previous studies support adjustment for stutter only in samples where the stutter is the result of a shared allele. As demonstrated in 20 patient samples, use of stutter cut-off values for markers with shared alleles reduces the bias in estimating chimerism and may improve the accuracy of estimated chimerism. This is the first study to systematically evaluate the contribution of stutter at all different levels of chimerism. Based on results from this study, the relative contribution of stutter to estimated chimerism appears to be significant

at low levels (5%) of chimerism and its contribution becomes less important at higher levels of chimerism.

Using median values for laboratory-defined stutter cut-off values resulted in substantially lower cut-off values for all markers as compared with the manufacturer-defined cut-off values. The improved accuracy of estimated chimerism and ability to consistently detect the 5% chimerism in all analytical runs using the laboratory-defined marker-specific stutter cut-off values as compared with manufacturer-defined cut-off values demonstrates the importance of determining laboratory-specific cut-off values for all STR markers. Many laboratories commonly use a 15% stutter cut-off value for all markers to uniquely identify the presence of a particular individual (9). Use of a uniform high stutter cut-off value across all markers is a conservative approach, and in the context of HCT, will result in decreased sensitivity to identify low-level chimerism. This study also confirmed results from a previous study that markers labeled with the NED dye had lower median stutter as compared with markers labeled with the JOE or FAM dyes (7). This supports the hypothesis that the repeat structure in the markers labeled with the JOE/FAM dyes may be associated with higher stutter.

A major practical limitation to adjustment of stutter, as described in this study, is the extensive post-PCR data manipulation. Advances in laboratory information management systems allow us to automate much of the post-PCR data analyses. In addition to increased efficiency, automation also allows us to decrease errors in data processing and will be a critical component of implementing these methods in the clinical molecular diagnostics laboratory. Another limitation of the current approach is that adjustment for stutter does not take into account allele size. It has been previously demonstrated that the size of the repeat alleles is positively correlated with stutter percentage (7,10). Although using allele sizes to determine a cut-off range for the stutter may give more repeat specific cut-off values, this approach would work only for common

alleles and it would be difficult to obtain reliable estimates of specific allele cut-offs for the unusual alleles. Finally, even though adjustment for stutter from a shared allele reduces the bias in estimating chimerism, it may result in inability to detect low-level chimerism in some instances. Hence, caution should be exercised when adjusting for stutter in patient samples with low levels of chimerism and, wherever possible, the results should be confirmed using other markers.

This study provides a framework for evaluating use of markers with overlapping stutter for determining chimerism after HCT. Though markers with overlapping stutters should not be used as a primary marker to determine the presence of individual components, these markers, with appropriate adjustment for stutter, can be used in combination with other markers to provide a more accurate estimate of chimerism in post-HCT samples.

ACKNOWLEDGMENTS

We acknowledge technical support of Teresa Kemmer, Luan Nguyen, Matt Schomaker, Laurie Pohlman, Helen Purvey, Katherine Rhakola, Stacy Hendry, and Kylene Karnuth in performing the assays described in this manuscript.

REFERENCES

1. Muller-Berat N, Lion T. Chimerism and transplant-related diagnostics. *Leukemia* 2006;20:1358–1360.
2. Nollet F, Billiet J, Selleslag D, Criel A. Standardisation of multiplex fluorescent short tandem repeat analysis for chimerism testing. *Bone Marrow Transplant* 2001;28:511–518.
3. Bennett P. Demystified ... microsatellites. *Mol Pathol* 2000;53:177–183.
4. Schichman SA, Suess P, Vertino AM, Gray PS. Comparison of short tandem repeat and variable number tandem repeat genetic markers for quantitative determination of allogeneic bone marrow transplant engraftment. *Bone Marrow Transplant* 2002;29:243–248.
5. Hauge XY, Litt M. A study of the origin of “shadow bands” seen when typing dinucleotide repeat polymorphisms by the PCR. *Hum Mol Genet* 1993;2:411–415.
6. Walsh PS, Fildes NJ, Reynolds R. Sequence analysis and characterization of stutter products at the tetranucleotide repeat locus vWA. *Nucleic Acids Res* 1996;24:2807–2812.
7. Leclair B, Fregeau CJ, Bowen KL, Fourney RM. Systematic analysis of stutter percentages and allele peak height and peak area ratios at heterozygous STR loci for forensic casework and database samples. *J Forensic Sci* 2004;49:968–980.
8. Chen DP, Tseng CP, Tsai SH, Wu TL, Chang PY, Sun CF. Systematic analysis of stutters to enhance the accuracy of chimerism testing. *Ann Clin Lab Sci* 2008;38:264–272.
9. Butler JM. *Forensic DNA Typing: Biology, Technology, and Genetics of STR Markers*. Boston: Elsevier Academic Press, 2005. 632 p.
10. Holt CL, Buoncristiani M, Wallin JM, Nguyen T, Lazaruk KD, Walsh PS. TWGDAM validation of AmpFISTR PCR amplification kits for forensic DNA casework. *J Forensic Sci* 2002;47:66–96.