Blood
Disorders & Transfusion

Nakorchevsky et al., J Hematol Thromb Dis 2016, 7:1
http://dx.doi.org/10.4172/2155-9864.1000340

Research Article Open Access

Detection of Low Level Mixed Chimerism Using High Throughput SNP Genotyping

Aleksey A. Nakorchevsky[1], Eunice Flores[1], Li Xiangyang[2], Tao Hong[2] and Anders O.H. Nygren

[1]*Agena Bioscience, San Diego, CA, USA*

[2]*Hackensack University Medical Center, Hackensack NJ, USA*

Abstract

Recipients of allogeneic bone marrow transplants (BMT) or stem cell transplants (SCT) require clinical monitoring to allow for early diagnosis of such post-transplant adverse effects as rejection, graft vs. host disease (GVHD), or a malignancy relapse. Triaging of the transplant recipients in clinical settings is achieved by monitoring the Minimal Residual Disease (MRD) and measuring the amount of mixed chimerism in peripheral blood lymphocytes (PBL). While MRD monitoring involves detection of the malignancy-specific markers, measuring the extent of mixed chimerism can be achieved via general PCR-based methods. We have developed a SNP genotyping method to detect low levels of mixed chimerism in PBL and genomic DNA. Sensitivity is achieved by measuring a cumulative skew in genotyping data across a cohort of 92 independent SNP markers. This method showed a sensitivity of 0.98 and a specificity of 0.90 for 10%, 5%, and 2% mixed chimerism samples. The overall specificity of the method is 0.98 and the accuracy is 0.95. The results show 100% concordance with the STR data for a set of clinical samples. The advantage of this method compared to already established methodologies is that it does not require disease-specific markers and can be multiplexed. The method and the analysis software can also be used with other genotyping and sequencing technologies.

Keywords: SNP-based method; Mixed chimerism; Therapeutic response; Hematologic relapse

Introduction

Patients diagnosed with blood and bone marrow malignancies typically receive allogeneic bone marrow transplant (BMT) or stem cell transplant (SCT) treatment following chemotherapy. Due to a high rate of relapse post-allogeneic transplant [1], it is important to monitor the status of post-transplant patients to allow for early diagnosis of such adverse effects as transplant rejection, graft vs. host disease (GVHD), or a malignancy relapse. Clinical follow up of the allograft recipients requires long-term monitoring of either the malignancy-specific markers or the residual hematopoiesis of the recipient cells (mixed chimerism, MC). Minimal residual disease (MRD [1,2]) monitoring involves detection and quantification of the malignancy-specific markers, while measuring the extent of mixed chimerism [3] requires measuring the fraction of recipient cells in peripheral blood samples. MRD approaches require monitoring the malignant clone via molecular (PCR-based) or immunophenotypic methods [4]. For example, the gene-specific markers for acute lymphoblastic leukemia (ALL) include TCR and Ig-gene rearrangements and BCR/ABL gene fusion products [4], while mutations in the NPM1 gene can be used as the markers in CML [5].

Mixed chimerism (MC) is defined as presence of an additional genotype in the peripheral blood cells of the allograft recipient. This is attributed to the hematopoiesis of the recipient's native CD34+ derived cells despite the cytoreductive treatment prior to the allograft transplant. Several studies have shown strong positive correlation between the extent of mixed chimerism and the likelihood of patient hematologic relapse [4,6-9]. The magnitude of mixed chimerism is frequently measured using PCR-based methods [10,11] as a ratio of the recipient genotype signal to that of the donor. Some of the methods employed to estimate the degree of mixed chimerism include variable number tandem repeat (VNTR) PCR [3,12], short tandem repeat (STR) PCR [13,14], amelogenin marker PCR [15], single nucleotide polymorphism (SNP) marker PCR [16] and cell sorting analysis (FACS) [17].

The main difference between MRD and MC molecular detection methods is that MRD methods require disease-specific molecular probes while the MC methods identify mixed chimerism using universal non allele-specific genotyping PCR-techniques such as STRs, VNTRs and SNPs. For example the mutations in the NPM1 gene can be interrogated via quantitative or nested PCR to monitor patient samples with "normal karyotype AML" via MRD. However, simultaneous presence of the internal tandem duplication mutations in FLT3 indicates higher rates of chemotherapy failure and allograft hematologic relapse [5]. MRD detection molecular targets are disease-, subgroup-, and often patient-specific, which complicate creating a universal methodology. MRD does offer higher sensitivity than MC methods, ranging around $1e^{-3}$-to-$1e^{-4}$ (1 molecule in 10,000) [17-21]. MC methods, on the other hand, use genotyping techniques that can be applied to various blood cancers in a manner that is independent of the mutation makeup and the karyotype of the malignancy. The early MC detection methods based on STRs and VNTRs claimed around 3-5% mixed chimerism sensitivity [3,13,14]. Recent SNP-based methods coupled with qPCR [16,22] report quantification of the target template as low as 0.1%. For reviews of MRD and MC detection methods see [4,5,23].

We have developed a high-throughput SNP-based mixed chimerism detection method that is highly multiplexed and does not involve allele-specific PCR. The 92 SNP ChimericID Panel was developed for the MALDI-TOF-based MassARRAY® System with iPLEX® technology by Agena Bioscience™ [24-27] for the purpose of mixed chimerism detection in PBLs and genomic DNA. The panel consists of 92 independent SNPs with minor allele frequency (MAF) of 0.45-to-0.55. Mixed chimerism detection is achieved by detecting a cumulative skew

***Corresponding author:** Aleksey A. Nakorchevsky, Ph.D., M.Sc., Sr. Scientist Bioinformatics, Agena Bioscience, Inc.4755 Eastgate MallSan Diego, CA 92121, USA, Tel: 858-882-2693; E-mail: aleksey.nakorchevsky@agenabio.com

Received January 23, 2016; **Accepted** February 22, 2016; **Published** February 25, 2016

Citation: Nakorchevsky AA, Flores E, Xiangyang L, Hong T, Nygren AOH (2016) Detection of Low Level Mixed Chimerism Using High Throughput SNP Genotyping. J Blood Disord Transfus 6: 340. doi:10.4172/2155-9864.1000340

in genotyping data across a cohort of 92 markers. Detection of as low as 2% mixed chimerism was achieved with a sensitivity of 0.98 and a specificity of 0.90. The advantage of the method compared to already established methodologies is that it does not require disease-specific markers and can be multiplexed to process samples in parallel (4, 12, or 48 samples per 24, 96, or 384 microtiter plate format, respectively). The method and the analysis software can also be used with other genotyping and sequencing technologies and applied to general tasks of sample mixture identification.

Methods

Biological samples

The panel design and the serial dilution experiments were carried out with DNA from Coriell HapMap CEU population 96-samples plate (Coriell HAPMAPPT01) and an in-house biobank of genomic DNA from four self-declared ethnic populations (Caucasian, Asian, African American, and Hispanic). Concentration and mixture composition of the samples was verified with a NanoDrop spectrophotometer (ThermoScientific). The panel was also tested using clinical PBL samples from a cohort of bone marrow transplant recipients (Hackensack UMC, NJ).

Panel design

Panel SNP candidates were selected using the International HapMap Project [28] data retrieval utility (BioMart). We selected a cohort of SNPs with average MAF within the range of 0.45 to 0.55 across the following major HapMap populations: ASW, CEU, CHB, GIH, JPT, MEX. The cohort was reduced to 228 SNPs to include only A/T and C/T transitions. This decision was made to take advantage of the strengths of the MassARRAY System and iPLEX chemistry, rather than the biological significance of the allele transitions. The SNAP utility from Broad Institute (http://www.broadinstitute.org/mpg/snap/) was used to verify the absence of linkage disequilibrium among the panel SNP markers from the SNP data available from the International HapMap Project [29]. We used the Agena Bioscience Assay Design Suite 1.0 (Agena Bioscience ADS 1.0) to design the multiplexed iPLEX assays. The final panel contains 92 assays in 8 multiplexes (Table 1).

Biochemistry and data acquisition

10-20 ng of the total DNA per multiplex (80-160 ng of DNA per sample) was used for locus-specific PCR amplification (45 cycles) with the PCR Accessory and Enzyme Set (Agena Bioscience) in both 96- or 384-well format. Following PCR amplification of the panel marker targets, the unincorporated dNTPs were dephosphorylated with SAP (Agena Bioscience) and the assay probes were extended into SNP sites by single-nucleotide extension with acyclo-NTP termination (iPLEX Pro Reagent Kit, 200 cycles of the extend reaction, Agena Bioscience). The contents of the PCR/extend reactions were desalted with strong cation-exchange resin and transferred onto a SpectroChip® Array (SpectroCHIP Array and Resin Kit, Agena Bioscience) for MALDI-TOF analysis. Data were acquired with the MassARRAY Analyzer 4 (Agena Bioscience) in genotyping plus area mode. The signal processing, peak annotation, and genotype calling were done with TyperAnalyzer 4.0 (Agena Bioscience).

MassARRAY iPLEX data

The iPLEX extension products are detected with MALDI-TOF MassARRAY Analyzer. Figure 1 shows examples of the iPLEX data; Figure 1A is a multiplexed spectrum and Figure 1B and 1C show homozygous and heterozygous SNP data. SNP allele frequencies and

rs1035271	rs1481847	rs4387937	rs7102303
rs1044910	rs153749	rs441460	rs738940
rs10454068	rs1854853	rs4586881	rs754666
rs10744034	rs1893023	rs4588273	rs7652997
rs10771010	rs1893673	rs4722897	rs77774941
rs10898438	rs1958693	rs4751376	rs7779384
rs10956179	rs1971391	rs4810570	rs7842255
rs1105359	rs2026263	rs643502	rs7848571
rs11136962	rs217297	rs643864	rs7903919
rs11670999	rs2202828	rs6442180	rs7967526
rs11671916	rs2236067	rs6455790	rs8037351
rs11989001	rs2255054	rs652539	rs838640
rs11989433	rs2256763	rs6657751	rs924758
rs1206523	rs2358996	rs6684496	rs9291849
rs12086717	rs246718	rs6691155	Rs9372042
rs12313812	rs2548467	rs66939853	rs945246
rs12318959	rs2548468	rs6782792	rs9458428
rs12872438	rs2561243	rs6811404	rs9552210
rs12919933	rs2615519	rs685553	rs962197
rs13140054	rs284661	rs6927989	rs9729034
rs1387566	rs3745290	rs6938076	rs9862216
rs1446940	rs4241782	rs6994603	rs9901757
rs1451509	rs4322101	rs7070511	Rs9927268

Table 1: ChimericID panel SNP markers.

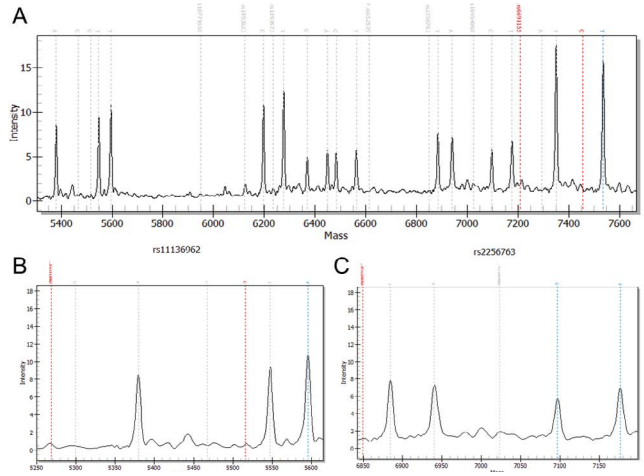

Figure 1: Examples of MassARRAY iPLEX data. Figure A shows a multiplexed (10 assays) MALDI-TOF spectrum of ChimericID Panel data. Figures B and C shown the assays that are homozygous (B) and heterozygous (C). In B and C the SNP-related peaks are highlighted with blue and red dashed lines.

the ensuing genotypes are derived from intensities of the ion peaks for that SNP. The numeric genotype for each SNP is expressed as follows: $\frac{Int_{low\ mass}}{(Int_{low\ mass} + Int_{high\ mass})}$, where $Int_{low\ mass}$ is the peak intensity of the low mass allele and $Int_{high\ mass}$ is the peak intensity of the high mass allele for the assays with two allele products. Numeric genotype values range from 0 to 1. The MAF is derived as the following ratio: $\frac{Int_{minor}}{(Int_{major} + Int_{minor})}$, where Int_{major} is the peak intensity of the major allele and int_{minor} is the peak intensity of the minor allele for the assays with two allele products.

Chimerism detection algorithm

Chimerism is detected by measuring allele imbalance (assay skew)

in a subset of panel assays that are informative (SNPs that have different genotypes in donor and recipient). Mixed chimerism is determined by comparing the MAF and the assay skew of the unknown samples to the data from known non-chimeric samples (universal approach). Historical MAF baseline data are used for the comparisons and the z-score is calculated for each panel SNP marker (absolute z-score value is calculated using the mean and the standard deviation of the baseline data). The per-assay z-score represents the distance of that sample's assay from non-chimeric baseline. The top n (50 by default) z-scores are averaged and interpreted as the overall sample z-score. In addition to z-scores that represent a confidence measure that the sample is chimeric, the sample MAF of the informative type I assays is averaged and reported as the percentage of mixed chimerism. The informative type I assays are the SNP markers that are homozygous for different genotypes for both the donor and the recipient.

Software

The analysis software was developed using Perl and R languages and was integrated with Agena Bioscience's TyperAnalyzer 4.0 as an executable plugin.

Results

Mixed chimerism is highly prognostic of hematologic relapse. Of the molecular methods that measure MC the majority involve microsatellite analysis of STRs and VNTRs, with only a few SNP-based methods [16,22]. The ChimericID Panel contains 92 SNP markers that are listed in Table 1. The markers, selected as described in Methods from International HapMap Project data [28], are bi-allelic, independent and not in linkage across six major HapMap populations.

ChimericID information content

In seeking to establish the information content of the panel for the purposes of discriminating any two unique samples, we estimate the discriminating power of the panel in the following fashion. Each bi-allelic panel SNP has the following three classes of outcomes: *AA* (homozygous dominant), *Aa* (heterozygous) and *aa* (homozygous recessive). For the search space T, x_{AA}, x_{Aa}, x_{aa} are the number of observations of each class with the fractional probability of belonging to a particular class of $p_i = \chi_i / T$. Given that the total number of pairwise comparisons between the samples of a search space T is $T(T-1)/2$, the probability of a match between any two unique samples is $\frac{\sum x_i (x_i - 1)}{T(T-1)}$ or $P = \frac{\sum p_i^2 - 1/T}{1 - 1/T}$. Approximating the T to be large: $P = \sum p_i^2$ where p_i are probabilities of AA, Aa, and aa. Therefore the probability of a random match between any two samples using a single SNP marker is:

$$P = \sum p_i^2 = p_{AA}^2 + p_{Aa}^2 + p_{aa}^2,$$

while the probability of a random match between any two samples using a cohort of 92 SNPs is:

$$P = \prod_{j=1}^{92} P_j$$

Figure 2 shows the relationship between the number of SNP markers in the panel and the various outcomes of random matches (probability, expected value, and cumulative probability). For a 92 SNP panel the probability of a random match between any two samples is 6.5×10^{-40} (Figure 2 blue trace), while the expected value (e-value, the number of the actual random matches) considering the current size of the world population (order of 3×10^9) is 1.9×10^{-31} (Figure 2 red trace). The probability of a random match for all possible pairwise interactions among the world population is 5.7×10^{-22}.

ChimericID informative assays

The principle of mixed chimerism detection with the ChimericID Panel is based on detecting small levels of allele imbalance (assay skew) in a subset of informative assays. For a hypothetical mixed chimerism scenario where the recipient R receives the allograft from the donor D, the informative SNPs are the markers with different genotypes in D and R. We refer to the SNPs where the D and the R are both homozygous (and have different genotypes) as informative type I and the SNPs where either the D or the R is homozygous and the other is heterozygous as informative type II. Assuming that the panel SNPs are independent, for any two unrelated individuals the probability of the informative type I scenario is 0.25 and that of the informative type II scenario is 0.375 per panel SNP. We used binomial distribution to calculate that for a 100 SNP panel there will be at least 12 to 39 informative type I SNPs and 23 to 51 informative type II SNPs at a 0.99 confidence level. The most likely number of informative type I SNPs is 24 and of informative type II SNPs is 38. Therefore, for the 92 SNP panel we expect somewhere between 35 and 90 informative SNPs, with 57 being the most likely outcome. Figure 3 shows the ChimericID Panel informative SNP data collected using unrelated HapMap samples (CEU population) combined in 1:1 ratio (50% mixed chimerism). We have observed an average of 55 informative markers in 48 mixed HapMap samples (compared to the 57 predicted). A high correlation between the theoretically derived and the experimentally determined number of informative SNPs confirms that the panel SNPs are not in linkage, with a MAF of approximately 0.5.

Sample mixtures

Samples with mixed chimerism are detected and quantified by comparing the numeric genotype of the SNP from an unknown sample to the numeric genotype of the same SNP from non-chimeric "baseline" samples. Historical baseline data that include all available genotypes for each SNP are used in calculating per-assay z-scores for each unknown sample. The top n highest z-scores are averaged and reported as the sample z-score. The average MAF is also calculated from the assays used in z-score calculations. Figure 4 shows the per-marker data for three synthetic chimeric samples, the 10%, 5%, and 2% mixed chimerism samples. The red trace shows the per-marker z-scores of samples with mixed chimerism (10%, 5%, and 2% in Figure 4A-4C) while the blue and green traces show the z-score of the samples used

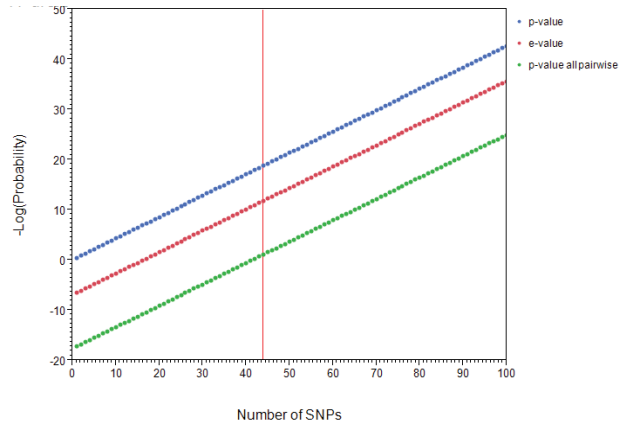

Figure 2: Relationship between the number of SNPs in the panel and the probability of a random positive match between any two unique samples, expressed as a p-value (blue trace), the expected number of false positives for the population of an entire planet (red trace), and the probability of a random positive match considering all pairwise comparisons for the world population (green trace).

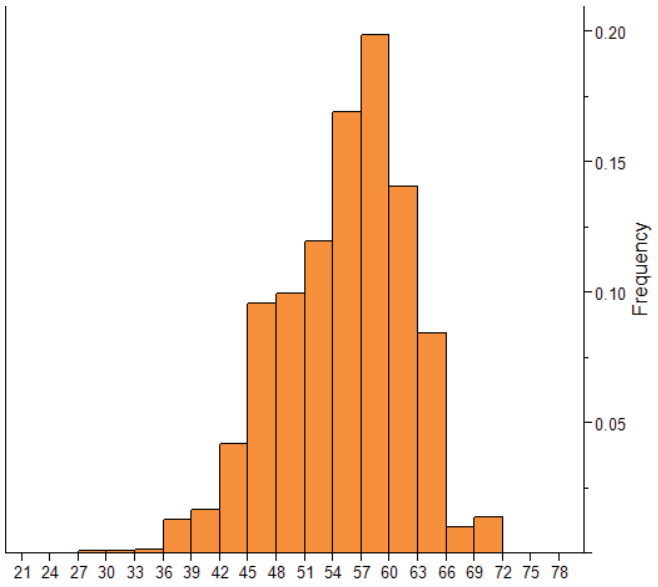

Figure 3: Distribution of the number of informative SNPs for 44 chimeric (50% chimerism) HapMap CEU samples.

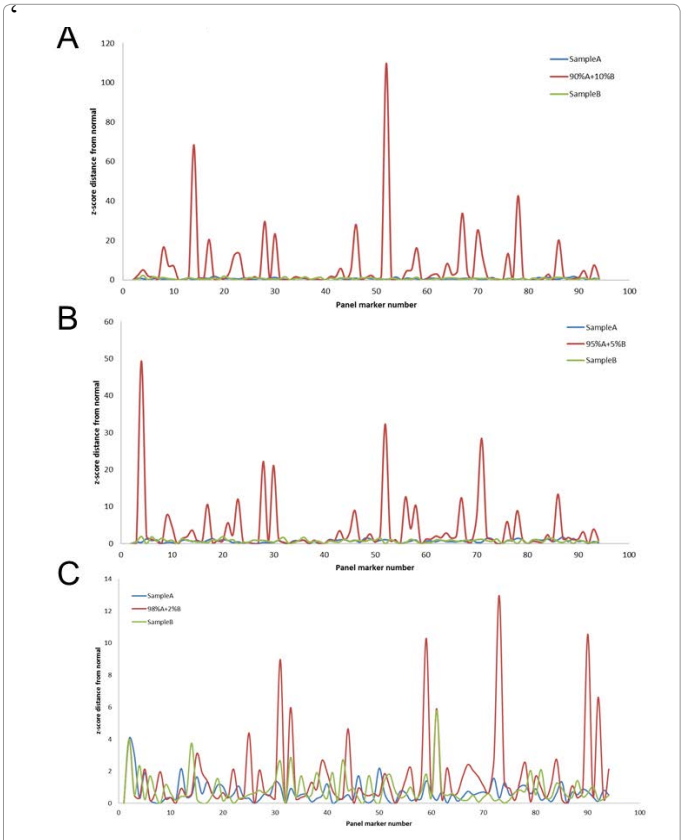

Figure 4: Per marker z-score (y-axis) of non-chimeric samples (Sample A-blue trace, Sample B-green trace) and chimeric samples (red trace), and three levels of mixed chimerism: 10% (A), 5% (B), and 2% (C). Panel markers are shown on the x-axis.

to create the chimeric mix. It is evident that the informative assays in mixed chimerism samples have a significantly higher skew from the normal baseline, which results in large "spikes" of the z-scores calculated using the baseline data. The magnitude of z-score deviations for the informative assays decreases as the percent chimerism decreases (Figure 4A-4C). The top *n* (50 by default) SNPs are used when determining the average z-score for each sample. The per-sample average z-scores are summarized in the following Table 2.

The z-scores for the Sample A + B for 2%, 5%, and 10% levels of mixed chimerism are significantly higher than those for non-chimeric samples (Sample A and Sample B).

Population study

The panel was further tested using samples from various ethnic populations. The samples were combined to create mixed chimerism of 10%, 5%, and 2%. The following table lists the composition of the sample sets used in development and benchmarking of the ChimericID Panel (Table 3).

Figure 5 shows performance of the method for the samples listed. Data points in Figure 5A show overall z-scores (see Methods for definition) for all samples of different amounts of mixed chimerism (color scheme). The average z-scores increase with the increasing amount of mixed chimerism. This allows differentiating chimeric samples from normal samples based on the average z-score. Figure 5B shows the receiver operating characteristic (ROC) plot of the panel performance. The optimal separation between the non-chimeric samples (purple data points) and chimeric samples is achieved at a z-score of 0.86. The cutoff yields an overall TPR of 0.98 and an FPR of 0.063. The total area under the curve of the panel ROC plot is 0.976. Below is the confusion matrix that lists the classification outcomes (Table 4).

Chimeric+ and Chimeric- represent the actual sample categories of either chimeric or non-chimeric samples and Software+ and Software- indicate the software classification outcome as chimeric or non-chimeric. The method has a sensitivity of 0.98 and a specificity of

	2%	5%	10%
Sample A, major component	0.02	0.09	0.08
Sample A+B	0.88	3.28	5.17
Sample B, minor component	0.17	0.01	0.00

Table 2: Average z-scores per-sample.

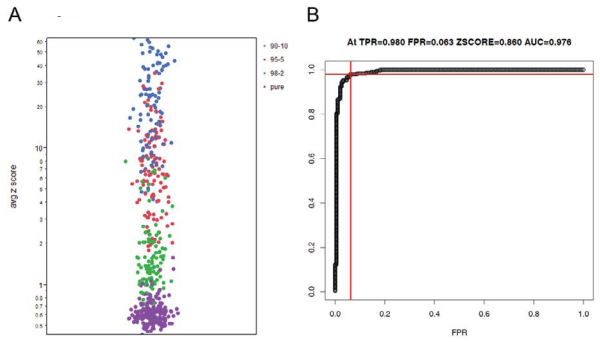

Figure 5: (A) Average z-score of chimeric samples of different percentages (90-10-blue, 95-5-red, 98-2-green, non-chimeric – purple). (B) The ROC curve of the overall panel performance (TRP – true positive rate y-axis, FPR – false positive rate x-axis). Red horizontal line indicates TPR of 0.98 and the vertical line shows the corresponding FPR (0.063).

	Chimeric (10%, 5%, 2%)	Normal
HapMap CEU cohort	48 (16,16,16)	32
Ethnic pop (AF, AS, HI)	144 (48,48,48)	96
HapMap CEU related	108 (36,36,36)	64

Table 3: Population study.

	Chimeric+	Chimeric-
Software+	TP (295)	FP (20)
Software-	FN (5)	TN (172)

Table 4: List of classification outcomes.

0.90 for the 10%, 5%, and 2% mixed chimerism samples. The overall specificity of the method is 0.98 and the accuracy is 0.95. The HapMap CEU related cohort represents mixed samples derived from parent-child trios. In theory, the expected number of informative assays will reduce by a factor of 2 in parent-child trio samples. The reduced number of SNPs will result in a "weakened" signal from related chimeric samples. However, even for the HapMap CEU related cohort the sensitivity and specificity values are 0.95 and 0.90 respectively.

Panel content and performance

We have demonstrated in Figure 3 that the number of informative SNPs ranges from 28 to 72 for a chimeric sample. The impact of reducing the number of panel assays and, therefore, the number of informative markers is shown in Figure 6. Reducing the panel content to 2 multiplexes (23 assays, Figure 6A) or 4 multiplexes (47 assays, Figure 6B) results in reduced classification performance compared to the full complement of panel markers (92 assays, Figure 6C). The effect of reduced panel content on method performance is due to the decreasing number of informative SNPs in a mixture of two samples. As mentioned previously a mixture of any two samples would have a stochastic set of SNPs that are informative. Since chimerism is determined by considering the top "outlier" SNPs and averaging the pre-assay z-scores and the minor allele contributions, reduction in the total number of SNPs decreases the number of informative SNPs, and subsequently, the observed per-sample z-score and the minor allele frequency.

Clinical samples

Data for 12 clinical samples were collected and tested at the Hackensack University Medical Center according to the standard iPLEX protocol and the results were analyzed using ChimericID Panel software. Table 5 lists the analysis outcomes for the samples. The two sets of samples, ID101 and ID102, represent two donor-recipient sets. Columns 2 and 3 describe the samples. Each sample set (ID101 and ID102) contains donor and recipient pre-transplantation samples (expected non-chimeric) and several post-transplantation samples that represent different cell populations (PBL, CD3, and CD15 enrichments) collected at various times after the bone marrow transplantation.

Having non-chimeric pure donor and recipient samples for each instance of transplantation allows for a simplified analysis approach, the Known-Trio approach. The Universal approach (discussed in the Methods section and used extensively up to this point) does not make assumptions as to which markers are informative and uses a set of the most "significant" outlier markers to determine the per-sample z-score and the MAF. The Known-Trio approach is a simplified approach where the informative markers are determined from the pure donor and recipient samples and the MAF and the chimerism percentage are calculated directly from these markers. The blue column in Table 5 shows the Known-Trio approach results, the green columns show the Universal approach results, and the orange column shows the consensus ChimericID classification status (YES for chimeric). The last two columns in red show the results of the independent analysis of the samples with orthogonal technology (STR interrogation with qPCR). It is evident that the ChimericID Universal approach results (green columns) confirm non-chimeric status of the pure donor and recipient samples (z-score of less than 1 and the percentage of chimerism under 1%). The ChimericID mixed-chimerism results for the rest of the samples correlate very strongly with both methods of calculating mixed-chimerism (Universal and Known-Trio) as well as with the STR PCR method. The clinical status as determined by the ChimericID Panel is identical to that of the STR PCR approach.

The ID101 set contains samples collected at 6 and 8 weeks post-transplant. The results in the 8-week set of samples show that PBLs have lower mixed chimerism compared to the CD3-enriched fraction. This is probably due to the fact that PBL samples contain many different cell types that respond differently to the post-transplantation immunosuppressant treatment. For example, the CD15 fraction in both sample sets contains very low levels of mixed chimerism compared to the PBL and CD3 fractions. This could be due to the fact that the patients were receiving granulocyte-macrophage colony-stimulating factor (GM-CSF) post transplantation, which increased production of the donor-contributed cells.

Discussion and Conclusion

Close monitoring of allogeneic transplant recipients is required for early detection of hematologic relapse. Early studies of MC were conducted using STR analysis supplemented with gender profiling [3,13,14], followed by electrophoresis-based sequencing. Further work showed that combining leukocyte CD34+ subset enrichment via FACS followed by STR analysis improved the method's sensitivity to be on par with MRD detection [17]. One of the first SNP-based mixed chimerism detection methods, published in 2003, involved monitoring of 14 SNPs via pyrosequencing. The authors reported an average of 1 informative SNP for a chimeric sample in their study and a sensitivity level of 5% mixed chimerism. More sensitive qPCR-based and Taqman RT-PCR SNP methods were able to take advantage of allele-specific amplification and achieve a 0.1% detection limit [16,22]. The allele-specific PCR enrichment methods require knowledge of the donor and the recipient genotypes for the SNP markers and are usually handicapped by low sample throughput and high cost due to the limited multiplexing capabilities.

Here we described a SNP-based mixed chimerism detection panel. We implemented the panel on Agnea Bioscience's MALDI-TOF MassARRAY System using iPLEX methodology. The targeted nature of the iPLEX method offers unprecedented specificity when analyzing SNPs due to the two-stage selection. The first stage is the locus-specific PCR amplification of the assay targets. Therefore, unlike universal PCR methods often employed in WGS experiments, the ChimericID Panel library contains only 92 targets. The second stage of the selection is the specificity of the assay probe that anneals upstream of the SNP site and extends a single nucleotide into the SNP site. Finally, the sensitivity of the MALDI-TOF mass spectrometer that uses ionized assay products for detection without secondary reporters such as fluorophores or light quenchers contributes to the high specificity and sensitivity of iPLEX chemistry on the MassARRAY System.

Early detection of residual hematopoiesis of the recipient CD34 cells require detection of mixed chimerism of 5% or lower. The sensitivity of

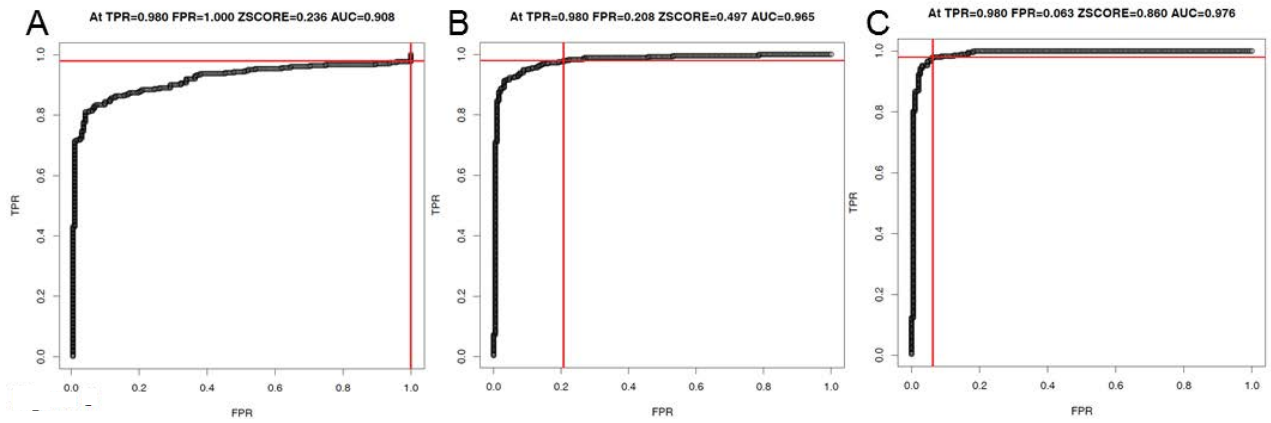

Figure 6: Relationship between panel content and performance. (A) 2 multiplexes (23 assays), (B) 4 multiplexes (47 assays), (C) Full panel of 8 multiplexes (92 assays).

Sample Information			Outcome					
Sample	Description	Sample Type	Known-Trio Approach (%)	Universal Approach		ChimericaID Status	Hacken Sack UMC (%)	Clinical Status
				Zscore	%			
ID101_D	Donor, Pre transplant	Pure	-	0.92	0.09	NO		NO
ID101_R	Recipient, Pre transplant	Pure	-	1.03	0.5	NO		NO
ID101_Post8_PBL	Recipient, Post transplant PBL	Possible chimerism	2.9%	12.72	2.66	YES	8	YES
ID101_Post8_CD3	Recipient, Post transplant T cell receptor	Possible chimerism	19.3%	89.12	16.25	YES	33	YES
ID101_Post8_CD15	Recipient, Post transplant neutrophil	Possible chimerism	0.06%	0.74	0.19	NO	0	NO
ID101_Post6_PBL	Recipient, Post transplant PBL	Possible chimerism	3.8%	56.41	3.5	YES	6	YES
ID101_Post6_CD15	Recipient, Post transplant neutrophil	Possible chimerism	0.88%	0.88	0.13	NO	0	NO
ID102_D	Donor, Pre transplant	Pure	-	0.86	0.29	NO		NO
ID102_R	Recipient, Pre transplant	Pure	-	0.97	0.25	NO		NO
ID102_Post13_PBL	Recipient, Post transplant PBL	Possible chimerism	3.15%	6.84	2.99	YES	3	YES
ID102_Post14_PBL	Recipient, Post transplant PBL	Possible chimerism	16.96%	80.97	16.4	YES	14	YES
ID102_Post15_PBL	Recipient, Post transplant PBL	Possible chimerism	9.22%	40.66	9.4	YES	16	YES

Table 5: ChimericID clinical samples.

an individual iPLEX assay for minor allele detection is 5% or better [29]. This is due to the limited dynamic range of the mass spectrometer, where the minor low intensity peak will not be detected if the ratio between the major and the minor assay products exceeds 50X. Therefore, considering that the minimum minor allele frequency that can be detected by a single informative iPLEX assay is only 5%, in order to achieve higher sensitivity we have implemented a method where the weak minor allele signal is amplified with many informative SNPs and the final mixed chimerism percentage is the average of the assays that are outliers compared to their performance for normal non-chimeric samples.

We have tested the panel using several sample populations. The samples include non-chimeric samples, chimeric samples constructed with a HapMap Caucasian population (HAPMAPPT01, Coriel) with unrelated and related samples (parent-child trios), the ethnic populations, and clinical PBL samples from patients that have received an allograft. The method has a sensitivity of 0.98 and a specificity of 0.90 for 10%, 5%, and 2% mixed chimerism samples. We also showed that the performance of the method decreases with a decreasing number

of assays (Figure 6); reducing the number of assays from 92 to 47 or 23 decreased the specificity and increased the false positive rate (FPR) from 0.063 to 0.208 and 1.0, respectively (all had a true positive rate of 0.98). Decreasing performance is due to the diminishing number of informative assays that contribute to the final MC significance and MAF.

Finally, we have tested the panel using clinical samples (Table 5). We have implemented an additional analysis option for cases where both the donor and the recipient pre-transplant samples are available. While the original Universal approach used the per-assay z-score distance to enrich for unconfirmed informative assays, the new approach (Known-Trio) learns informative assays directly from the donor and the recipient samples and calculates MAF and mixed-chimerism directly from those assays. Results from both analysis methods have a high degree of correlation. The results also correlate with the STR qPCR methodology used by the Hackensack UMC. The amount of mixed chimerism varies depending on the blood cell fraction type (PLB, CD3, or CD15). Peripheral blood lymphocytes are a mixture of different cell types that respond differently to post-transplant treatment and immunosuppressant regime, and so we expect a lower effective amount of mixed chimerism presented in PBLs.

The advantage of the SNP-based iPLEX method compared to already established methodologies is that it has higher sensitivity than STR-based methods, it does not require disease-specific markers or allele-specific PCR, and it can be multiplexed to increase sample throughput. The panel and the analysis software can also be used with other genotyping or sequencing technologies.

Author Contribution and Conflict of Interest Disclosure

Aleksey A. Nakorchevsky has developed the method, outlined the experiments, analyzed the data and prepared the manuscript. Eunice Flores has performed majority of the experiments. Li Xiangyang has supervised the experiments performed at Hackensack UMC. Tao Hong has served as a Principal Investigator at Hackensack UMC for this project. Anders O.H. Nygren has served as a Principal Investigator at Agena Bioscience.

The authors disclose that there are no actual or potential conflicts of interest that apply to the manuscript, including the authors' material interests or economic relationships. Aleksey Nakorchevsky, Eunice Flores and Anders Nygren are full time employees of Agena Bioscience and hold company stock options.

References

1. Knechtli CJ, Goulden NJ, Hancock JP, Grandage VL, Harris EL, et al. (1998) Minimal residual disease status before allogeneic bone marrow transplantation is an important determinant of successful outcome for children and adolescents with acute lymphoblastic leukemia. Blood 92: 4072-4079.
2. Knechtli CJ, Goulden NJ, Hancock JP, Harris EL, Garland RJ, et al. (1998) Minimal residual disease status as a predictor of relapse after allogeneic bone marrow transplantation for children with acute lymphoblastic leukaemia. Br J Haematol 102: 860-871.
3. Pindolia K, Janakiraman N, Kasten-Sportes C, Demanet C, Van Waeyenberge C, et al. (1999) Enhanced assessment of allogeneic bone marrow transplant engraftment using automated fluorescent-based typing. Bone Marrow Transplant 24: 1235-1241.
4. Bader P, Willasch A, Klingebiel T (2008) Monitoring of post-transplant remission of childhood malignancies: is there a standard?Bone Marrow Transplant 42 Suppl 2: S31-34.
5. Bacher U, Haferlach T, Fehse B, Schnittger S, Kröger N (2011) Minimal residual disease diagnostics and chimerism in the post-transplant period in acute myeloid leukemia. ScientificWorldJournal 11: 310-319.
6. Bader P, Kreyenberg H, Hoelle W, Dueckers G, Handgretinger R, et al. (2004) Increasing mixed chimerism is an important prognostic factor for unfavorable outcome in children with acute lymphoblastic leukemia after allogeneic stem-cell transplantation: possible role for pre-emptive immunotherapy? J Clin Oncol 22: 1696-1705.
7. Barrios M, Jiménez-Velasco A, Román-Gómez J, Madrigal ME, Castillejo JA, et al. (2003) Chimerism status is a useful predictor of relapse after allogeneic stem cell transplantation for acute leukemia. Haematologica 88: 801-810.
8. Gyger M, Baron C, Forest L, Lussier P, Lagacé F, et al. (1998) Quantitative assessment of hematopoietic chimerism after allogeneic bone marrow transplantation has predictive value for the occurrence of irreversible graft failure and graft-vs.-host disease. Exp Hematol 26: 426-434.
9. Won EJ, Kim HR, Kim HY, Kook H, Kim HJ, et al. (2013) Diagnostic and prognostic value of mitochondrial DNA minisatellites after stem cell transplantation. Biol Blood Marrow Transplant 19: 918-924.
10. Lawler M, Humphries P, McCann SR (1991) Evaluation of mixed chimerism by in vitro amplification of dinucleotide repeat sequences using the polymerase chain reaction. Blood 77: 2504-2514.
11. Ugozzoli L, Yam P, Petz LD, Ferrara GB, Champlin RE, et al. (1991) Amplification by the polymerase chain reaction of hypervariable regions of the human genome for evaluation of chimerism after bone marrow transplantation. Blood 77: 1607-1615.
12. Gaiger A, Mannhalter C, Hinterberger W, Haas O, Marosi C, et al. (1991) Detection of engraftment and mixed chimerism following bone marrow transplantation using PCR amplification of a highly variable region-variable number of tandem repeats (VNTR) in the von Willebrand factor gene. Ann Hematol 63: 227-228.
13. Chen DP, Tsao KC, Wang PN, Tseng CP, Sun CF (2002) Quantitative analysis of chimerism after allogeneic peripheral blood stem cell transplantation. Chang Gung Med J 25: 734-742.
14. Kreyenberg H, Hölle W, Möhrle S, Niethammer D, Bader P (2003) Quantitative analysis of chimerism after allogeneic stem cell transplantation by PCR amplification of microsatellite markers and capillary electrophoresis with fluorescence detection: the Tuebingen experience. Leukemia 17: 237-240.
15. Ghaffari SH, Chahardouli B, Gavamzadeh A, Alimoghaddam K (2008) Evaluation of hematopoietic chimerism following allogeneic peripheral blood stem cell transplantation with amelogenin marker. Arch Iran Med 11: 35-41.
16. Gineikiene E, Stoskus M, Griskevicius L (2009) Single nucleotide polymorphism-based system improves the applicability of quantitative PCR for chimerism monitoring. J Mol Diagn 11: 66-74.
17. Bornhauser M, Oelschlaegel U, Platzbecker U, Bug G, Lutterbeck K, et al. (2009) Monitoring of donor chimerism in sorted CD34+ peripheral blood cells allows the sensitive detection of imminent relapse after allogeneic stem cell transplantation. Haematologica 94: 1613-1617.
18. Rossi G, Carella AM, Minervini MM, Savino L, Fontana A, et al. (2013) Minimal residual disease after allogeneic stem cell transplant: a comparison among multiparametric flow cytometry, Wilms tumor 1 expression and chimerism status (Complete chimerism versus Low Level Mixed Chimerism) in acute leukemia. Leuk Lymphoma 54: 2660-2666.
19. Waterhouse M, Kunzmann R, Torres M, Bertz H, Finke J (2013) An internal validation approach and quality control on hematopoietic chimerism testing after allogeneic hematopoietic cell transplantation. Clin Chem Lab Med 51: 363-369.
20. Bernal T, Diez-Campelo M, Godoy V, Rojas S, Colado E, et al. (2014) Role of minimal residual disease and chimerism after reduced-intensity and myeloablative allo-transplantation in acute myeloid leukemia and high-risk myelodysplastic syndrome. Leuk Res 38: 551-556.
21. Terwey TH, Hemmati PG, Nagy M, Pfeifer H, Gökbuget N, et al. (2014) Comparison of chimerism and minimal residual disease monitoring for relapse prediction after allogeneic stem cell transplantation for adult acute lymphoblastic leukemia. Biol Blood Marrow Transplant 20: 1522-1529.
22. Qin XY, Li GX, Qin YZ, Wang Y, Wang FR, et al. (2011) Quantitative assessment of hematopoietic chimerism by quantitative real-time polymerase chain reaction of sequence polymorphism systems after hematopoietic stem cell transplantation. Chin Med J (Engl) 124: 2301-2308.
23. Khan F, Agarwal A, Agrawal S (2004) Significance of chimerism in hematopoietic stem cell transplantation: new variations on an old theme. Bone Marrow Transplant 34: 1-12.

24. Jurinke C, Denissenko MF, Oeth P, Ehrich M, van den Boom D, et al. (2005) A single nucleotide polymorphism based approach for the identification and characterization of gene expression modulation using MassARRAY. Mutat Res 573: 83-95.

25. Jurinke C, van den Boom D, Cantor CR, Köster H (2001) Automated genotyping using the DNA MassArray technology. Methods Mol Biol 170: 103-116.

26. Köster H, Tang K, Fu DJ, Braun A, van den Boom D, et al. (1996) A strategy for rapid and efficient DNA sequencing by mass spectrometry. Nat Biotechnol 14: 1123-1128.

27. Oeth P, del Mistro G, Marnellos G, Shi T, van den Boom D (2009) Qualitative and quantitative genotyping using single base primer extension coupled with matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MassARRAY). Methods Mol Biol 578: 307-343.

28. International HapMap Consortium (2003) The International HapMap Project. Nature 426: 789-796.

29. Jurinke C, Oeth P, van den Boom D (2004) MALDI-TOF mass spectrometry: a versatile tool for high-performance DNA analysis. Mol Biotechnol 26: 147-164.

Citation: Nakorchevsky AA, Flores E, Xiangyang L, Hong T, Nygren AOH (2016) Detection of Low Level Mixed Chimerism Using High Throughput SNP Genotyping. J Blood Disord Transfus 6: 340. doi:10.4172/2155-9864.1000340